

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 29, 2004, 10:13:38 ; Search time 118 Seconds
(without alignments)
7844.560 Million cell updates/sec

Title: US-09-712-338-1
Perfect score: 1668
Sequence: 1 atggctggctacgaattctt.....ccagtgtgtgtagcatag 1668

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/2/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PTCUS COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1662	99.6	1662	3	US-08-943-714-1
2	86.4	5.2	2503	1	US-09-640-305-3
3	86.4	5.2	2503	1	US-08-360-673-3
4	84	5.0	2002	1	US-08-309-341-3
5	84	5.0	2002	1	US-08-608-267-3
6	84	5.0	2002	1	US-08-608-452-3
7	84	5.0	2002	1	US-08-608-224-3
8	84	5.0	2002	2	US-08-967-149-3
9	76	4.6	2068	1	US-08-309-341-1
10	76	4.6	2068	1	US-08-608-267-1
11	76	4.6	2068	1	US-08-608-452-1
12	76	4.6	2068	1	US-08-608-224-1
13	76	4.6	2068	2	US-08-967-149-1
14	66.2	4.0	2632	3	US-08-899-324-32
15	66.2	4.0	2632	3	US-08-329-892B-32
16	56.4	3.4	1814	4	US-09-702-705-319
17	56.4	3.4	1814	4	US-09-736-457-319
18	56.4	3.4	1814	4	US-09-614-124B-319
19	56.4	3.4	1814	4	US-09-671-325-319
20	56.4	3.4	1814	4	US-09-589-184-319
21	47.6	2.9	390	3	US-09-197-649-7
22	47.6	2.9	1551	2	US-08-828-488-4
23	47.6	2.9	1551	4	US-09-299-689A-4
24	47.6	2.9	2076	4	US-09-907-794A-163
25	47.6	2.9	2076	4	US-09-905-125A-163
26	47.6	2.9	2076	4	US-09-902-775A-163
27	47.2	2.8	1670	2	US-08-828-488-2

28	47.2	2.8	1670	4	US-09-299-689A-2
29	45.4	2.7	1071	4	US-09-023-655-674
30	39.8	2.4	357	4	US-09-280-116-131
31	38.8	2.3	474	4	US-09-621-976-18033
32	38.6	2.3	447	4	US-09-252-991A-11967
33	38.6	2.3	630	4	US-09-252-991A-8154
34	38.6	2.3	774	4	US-09-252-991A-11893
35	38.6	2.3	777	4	US-09-252-991A-8222
36	38.6	2.3	1083	4	US-09-252-991A-11930
37	38.6	2.3	1239	4	US-09-252-991A-8276
38	38.6	2.3	1434	4	US-09-252-991A-8046
39	38.2	2.3	918	3	US-09-105-390-35
40	38.2	2.3	1008	3	US-09-105-390-51
41	38.2	2.3	1949	3	US-09-105-390-2
42	37.2	2.2	870	4	US-09-252-991A-10938
43	37.2	2.2	1785	4	US-09-252-991A-11178
44	37.2	2.2	2091	4	US-09-252-991A-11076
45	36.8	2.2	7012	4	US-09-221-017B-173

ALIGNMENTS

RESULT 1
US-08-943-714-1
; Sequence 1, Application US/08943714
; Patent No. 6187578
; GENERAL INFORMATION:
; APPLICANT: Blinkovsky, Alexander
; APPLICANT: Berka, Randy
; APPLICANT: Rev, Michael
; APPLICANT: Golightly, Elizabeth
; APPLICANT: Klotz, Alan
; APPLICANT: Mathisen, Thomas Erik
; APPLICANT: Dammann, Claus
; TITLE OF INVENTION: Carboxypeptidases And Nucleic Acids
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 61875780 No. 6187578disk of No. 6187578th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/943,714
FILING DATE: 03-OCT-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 4990.200-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1662 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 1..1662
OTHER INFORMATION:


```
;
; APPLICATION NUMBER: US/08/360,673
; FILING DATE: 06-FEB-1995
; APPLICATION NUMBER: WO PCT/FR93/00623
; FILING DATE: 23-JUN-1993
; APPLICATION NUMBER: FR 92/07785
; FILING DATE: 25-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, Julie K.
; REGISTRATION NUMBER: 38,619
; REFERENCE/DOCKET NUMBER: ST92040-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610)454-3839
; TELEFAX: (610)454-3808
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2503 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; ORGANISM: Kluyveromyces lactis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 387..1862
; OTHER INFORMATION: /product= "K. lactis protease C
; gene"
; /gene= "K1.PRC1"
;
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-640-305-3

Query Match          5.2%; Score 86.4; DB 1; Length 2503;
Best Local Similarity 57.4%; Pred. No. 6.8e-16;
Matches 156; Conservative 0; Mismatches 116; Indels 0; Gaps 0;

QY 199 GTCAAATCCTACTCTGGATATGTCGACACCTCTCCCGAGTCCCATACCTTTCTTCTGGTTC 258
Db      |||||
Db 624 GTGAACAATGTCGGGATATTAGATTACCGAGCTCAAAACACTCTTTTATTGGTTT 683
QY 259 TTGGAAGCCAGACATACCCAGAACTGCACCTATCATTGTGTTGAATGGTGGCCCT 318
Db      |||||
Db 684 TTGAGTCTAGAAATGACCCAGAAATGACCCAGTACTATGTTGTTAAACGGTGGTCT 743
QY 319 GGAAGCGATTCTTTGATCGGTCTCTTCGAAGAGTTGGGCCCTTCGCATGTCATTCGACT 378
Db      |||||
Db 744 GGCTGTTCTCTTTTCGTCGGTCTTTCTTTGAATTGGACCTTCTCTATAGGAGCTGAT 803
QY 379 TTTGATGACTACATCAACCCCTCATCTGGTGGAAACGAGGTCTCCAAATTTACTATTCTGTCC 438
Db      |||||
Db 804 TTGAACCCCAATTATACCCCTACTCTTGGAAATCCCAACGCTTCTGTGATATTCCTAGAT 863
QY 439 CAGCCATTGGAGTCGGCTTTTCATATAGTGA 470
Db      |||||
Db 864 CAGCCTGTGTTGGTGTCTCATACGGTGA 895

RESULT 3
US-08-360-673-3
; Sequence 3, Application US/08360673
; Patent No. 5679544
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; APPLICANT: Fournier, Alain
; APPLICANT: Yeh, Patrice
; TITLE OF INVENTION: MODIFIED KLUYVEROMYCES YEASTS, THEIR
; PREPARATION AND USE
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Rd. 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
```

```
;
; ZIP: 19002
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/360,673
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/FR93/00623
; FILING DATE: 23-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 92/07785
; FILING DATE: 25-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, Julie K.
; REGISTRATION NUMBER: 38,619
; REFERENCE/DOCKET NUMBER: ST92040-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610)454-3839
; TELEFAX: (610)454-3808
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2503 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; ORIGINAL SOURCE:
; ORGANISM: Kluyveromyces lactis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 387..1862
; OTHER INFORMATION: /product= "K. lactis protease C
; gene"
; OTHER INFORMATION: /gene= "K1.PRC1"
;
; US-08-360-673-3

Query Match          5.2%; Score 86.4; DB 1; Length 2503;
Best Local Similarity 57.4%; Pred. No. 6.8e-16;
Matches 156; Conservative 0; Mismatches 116; Indels 0; Gaps 0;

QY 199 GTCAAATCCTACTCTGGATATGTCGACACCTCTCCCGAGTCCCATACCTTTCTTCTGGTTC 258
Db      |||||
Db 624 GTGAACAATGTCGGGATATTAGATTACCGAGCTCAAAACACTCTTTTATTGGTTT 683
QY 259 TTGGAAGCCAGACATACCCAGAACTGCACCTATCATTGTGTTGAATGGTGGCCCT 318
Db      |||||
Db 684 TTGAGTCTAGAAATGACCCAGAAATGACCCAGTACTATGTTGTTAAACGGTGGTCT 743
QY 319 GGAAGCGATTCTTTGATCGGTCTCTTCGAAGAGTTGGGCCCTTCGCATGTCATTCGACT 378
Db      |||||
Db 744 GGCTGTTCTCTTTTCGTCGGTCTTTCTTTGAATTGGACCTTCTCTATAGGAGCTGAT 803
QY 379 TTTGATGACTACATCAACCCCTCATCTGGTGGAAACGAGGTCTCCAAATTTACTATTCTGTCC 438
Db      |||||
Db 804 TTGAACCCCAATTATACCCCTACTCTTGGAAATCCCAACGCTTCTGTGATATTCCTAGAT 863
QY 439 CAGCCATTGGAGTCGGCTTTTCATATAGTGA 470
Db      |||||
Db 864 CAGCCTGTGTTGGTGTCTCATACGGTGA 895

RESULT 4
US-08-309-341-3
; Sequence 3, Application US/08309341
; Patent No. 5594119
; GENERAL INFORMATION:
; APPLICANT: Yaver, Debbie Sue
; APPLICANT: Thompson, Sheryl Ann
; TITLE OF INVENTION: GENE ENCODING CARBOXYPEPTIDASE OF ASPERGILLUS NIGER
```


Patent No. 5693510
 GENERAL INFORMATION:
 APPLICANT: Yaver, Debbie Sue
 APPLICANT: Thompson, Sheryl Ann
 TITLE OF INVENTION: GENE ENCODING CARBOXYPEPTIDASE OF ASPERGILLUS NIGER
 NUMBER OF SEQUENCES: 4
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: No. 56935100 No. 5693510disk of No. 5693510th America, Inc.
 STREET: 405 Lexington Avenue, Suite 6400
 CITY: New York
 STATE: New York
 COUNTRY: U.S.A.
 ZIP: 10174-6401
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/608,452
 FILING DATE: 28-FEB-1996
 CLASSIFICATION: 536
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/309,341
 FILING DATE: 20-SEP-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Lowney, Karen A.
 REGISTRATION NUMBER: 31,274
 REFERENCE/DOCKET NUMBER: 4247.000-US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212 867 0123
 TELEFAX: 212 867 0298
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2002 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 ORIGINAL SOURCE:
 ORGANISM: Aspergillus niger
 FEATURE:
 NAME/KEY: intron
 LOCATION: 349..411
 FEATURE:
 NAME/KEY: CDS
 LOCATION: join (348..412)
 US-08-608-452-3

Query Match 5.0%; Score 84; DB 1; Length 2002;
 Best Local Similarity 55.5%; Pred. No. 3.3e-15;
 Matches 161; Conservative 0; Mismatches 129; Indels 0; Gaps 0;

QY	223	GACACCTCTCCGAGTCCCATACCTCTCTGTTCTTCGAAAGCCAGACATACCCAGAA	282
Db	617	GACAAACGAGAACGACAAAGCATCTCTCTGTTCTTCGAGTCTCGCAATGACCCGAG	676
QY	283	ACTGCACCTATCATTTGGTTGAATGGTGGCCCTGGAAGCGATTCTTTGATCGGTCTC	342
Db	677	AATGACCCCTGTGTTCTGTTGCTGAACGGTGGCCCTGGATGCTCTTCCCTCACCGGTCTT	736
QY	343	TTGGAAGAGTTGGCCCTTGCATGTCATTTTCGACTTTTTCATGACTACATCAACCTTCAC	402
Db	737	TTGATGAGCTCGGCCCTTAGCAGATCAACAGAGATCAGCCGGTCTCAACGACTAC	796
QY	403	TCGTGGAACGAGGTCTCCAAATTTACTATTCTGTCCAGGCAATGGGAGTCGGCTTTTCA	462
Db	797	GCTTGAACCTCAACCGGTCCGTGATCTTCTTGACCGAGCTGTCACGTCGGTTACTCT	856
QY	463	TATAGTGATACGGTTGATGGTTCATTAACCTGTAACTGGGGTCTGTCGA	512
Db	857	TACAGCAACTCTGCTGTGTCAGCAGACACCGTTGCTGTCGCAAGGACGTCTA	906

RESULT 7
 US-08-608-224-3
 Sequence 3, Application US/08608224
 Patent No. 5705376
 GENERAL INFORMATION:
 APPLICANT: Yaver, Debbie Sue
 APPLICANT: Thompson, Sheryl Ann
 TITLE OF INVENTION: GENE ENCODING CARBOXYPEPTIDASE OF ASPERGILLUS NIGER
 NUMBER OF SEQUENCES: 4
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: No. 57053760 No. 5705376disk of No. 5705376th America, Inc.
 STREET: 405 Lexington Avenue, Suite 6400
 CITY: New York
 STATE: New York
 COUNTRY: U.S.A.
 ZIP: 10174-6401
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/608,224
 FILING DATE: 28-FEB-1996
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/309,341
 FILING DATE: 20-SEP-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Lowney, Karen A.
 REGISTRATION NUMBER: 31,274
 REFERENCE/DOCKET NUMBER: 4247.000-US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212 867 0123
 TELEFAX: 212 867 0298
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2002 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 ORIGINAL SOURCE:
 ORGANISM: Aspergillus niger
 FEATURE:
 NAME/KEY: intron
 LOCATION: 349..411
 FEATURE:
 NAME/KEY: CDS
 LOCATION: join (348..412)
 US-08-608-224-3

Query Match 5.0%; Score 84; DB 1; Length 2002;
 Best Local Similarity 55.5%; Pred. No. 3.3e-15;
 Matches 161; Conservative 0; Mismatches 129; Indels 0; Gaps 0;

QY	223	GACACCTCTCCGAGTCCCATACCTCTCTGTTCTTCGAAAGCCAGACATACCCAGAA	282
Db	617	GACAAACGAGAACGACAAAGCATCTCTCTGTTCTTCGAGTCTCGCAATGACCCGAG	676
QY	283	ACTGCACCTATCATTTGGTTGAATGGTGGCCCTGGAAGCGATTCTTTGATCGGTCTC	342
Db	677	AATGACCCCTGTGTTCTGTTGCTGAACGGTGGCCCTGGATGCTCTTCCCTCACCGGTCTT	736
QY	343	TTGGAAGAGTTGGCCCTTGCATGTCATTTTCGACTTTTTCATGACTACATCAACCTTCAC	402
Db	737	TTGATGAGCTCGGCCCTTAGCAGATCAACAGAGATCAGCCGGTCTCAACGACTAC	796
QY	403	TCGTGGAACGAGGTCTCCAAATTTACTATTCTGTCCAGGCAATGGGAGTCGGCTTTTCA	462
Db	797	GCTTGAACCTCAACCGGTCCGTGATCTTCTTGACCGAGCTGTCACGTCGGTTACTCT	856

QY 373 TCGACTTTTGATGACTACATCAACCCCTCACTCGTGAACGAGGTCTTCCAAATTTACTATTC 432
Db 753 AAGAAGATCCAGCGGTCTTACAACTGACTACGTTTGAATCTCAACGCGTCCGTGATCTTC 812
QY 433 CTGTCCAGCCATTGGGAGTCGGCTTTTCATATAGTATGATACGTTTGATGGTCCATTAAAC 492
Db 813 CTGTACCAAGCCTGTCAATGTCTGTTACTCTTACAGTAAGTCTGTCTGTGTCAGCGACGCGTC 872
QY 493 CCTGTAACCTGGGTCGTGCA 512
Db 873 GCTGTGCAAGGACGTCTA 892

RESULT 10

US-08-608-267-1
; Sequence 1, Application US/08608267
; Patent No. 5688663
; GENERAL INFORMATION:
; APPLICANT: Yaver, Debbie Sue
; APPLICANT: Thompson, Sheryl Ann
; TITLE OF INVENTION: GENE ENCODING CARBOXYPEPTIDASE OF ASPERGILLUS NIGER
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 56886630 No. 56886630 disk of No. 56886630th America, Inc.
; STREET: 405 Lexington Avenue, Suite 6400
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/608,267
; FILING DATE: 28-FEB-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/309,341
; FILING DATE: 20-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowney, Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 4247.000-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 867 0123
; TELEFAX: 212 867 0298
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2068 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Aspergillus niger
; FEATURE:
; NAME/KEY: intron
; LOCATION: 572..632
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join (571..633)
US-08-608-267-1

Query Match 4.6%; Score 76; DB 1; Length 2068;
Best Local Similarity 55.8%; Pred. No. 8.9e-13;
Matches 145; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

QY 253 TGGTCTTCGAGCCACATACCCAGAACTGCACCTATCATATGCTGTTGAATGGT 312
Db 633 TGGTCTTCGAGTCTCGCAATGACCCCGAGAAATGATCCCGTTGTTCTGTGCTGCTGAACGGT 692

QY 313 GGCCCTGGAAGGATTCTTTGATCGGTCTCTTGAAGAGTTGGGCGCTTGGCATGFOAAT 372
Db 693 GGCCCTGGGTGCTCTTCCCTCACCAGTCTTTCATGAGCTTGGCCCTAGCAGCATCAAC 752
QY 373 TCGACTTTTGATGACTACATCAACCCCTCACTCGTGAACGAGGTCTTCCAAATTTACTATTC 432
Db 753 AAGAAGATCCAGCGGTCTTACAACTGACTACGTTTGAATCTCAACGCGTCCGTGATCTTC 812
QY 433 CTGTCCAGCCATTGGGAGTCGGCTTTTCATATAGTATGATACGTTTGATGGTCCATTAAAC 492
Db 813 CTGTACCAAGCCTGTCAATGTCTGTTACTCTTACAGTAAGTCTGTCTGTGTCAGCGACGCGTC 872
QY 493 CCTGTAACCTGGGTCGTGCA 512
Db 873 GCTGTGCAAGGACGTCTA 892

RESULT 11

US-08-608-452-1
; Sequence 1, Application US/08608452
; Patent No. 5693510
; GENERAL INFORMATION:
; APPLICANT: Yaver, Debbie Sue
; APPLICANT: Thompson, Sheryl Ann
; TITLE OF INVENTION: GENE ENCODING CARBOXYPEPTIDASE OF ASPERGILLUS NIGER
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 56935100 No. 56935100 disk of No. 56935100th America, Inc.
; STREET: 405 Lexington Avenue, Suite 6400
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/608,452
; FILING DATE: 28-FEB-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/309,341
; FILING DATE: 20-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowney, Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 4247.000-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 867 0123
; TELEFAX: 212 867 0298
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2068 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Aspergillus niger
; FEATURE:
; NAME/KEY: intron
; LOCATION: 572..632
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join (571..633)
US-08-608-452-1

Query Match 4.6%; Score 76; DB 1; Length 2068;
Best Local Similarity 55.8%; Pred. No. 8.9e-13;
Matches 145; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

253	QY	TGGTCTTCGAAGCCAGACATAACCCAGAACTGCACCTATACACTTGTGGTTGGAATGGT	312
633	Db	TGGTTCTTCAGACTCTCGCAATGACCCGAGAATGATCCCGTTCTCTGTGGCTGGAACGGT	692
313	QY	GGCCCTGGAAGCGATTCTTTTCATCGGTCTCTTCGAAGATTGGGCCCTCTGCCCATGTCAAT	372
693	Db	GGCCCTTGGGTGCTCTCCCTCACCGTCTCTTCATGGAGCTTGGCCCTAGCAGCATCAAC	752
373	QY	TGCGACTTTTGATGACTACATCAACCCCTCACTCGTGGAAAGAGGTTCTCCAAATTCATTATTC	432
753	Db	AAGAAGATCCAGCCGGTCTACAATGACTACCGCTTGGAACTCCAAACGGCTCCGTGATCTTC	812
433	QY	CTGTCCCGACCAATGGGAGTCGGCTTTTCATATGATGATACGGTTGATGGTCCATTAAC	492
813	Db	CTTGACCAACCTGTCAATGTCGGTTTCTCTCCTCAGTAACCTCTGCTGTACAGCACGGTTC	872
493	QY	CCTGTAACCTGGGGTCGTCTGA	512
873	Db	GCTGCTGGCAAGGAGTCTTA	892

RESULT, T 12

```

US-08-608-224-1
; Sequence 1, Application US/08608224
; Patent No. 5705376
; GENERAL INFORMATION:
; APPLICANT: Yaver, Debbie Sue
; APPLICANT: Thompson, Sheryl Ann
; TITLE OF INVENTION: GENE ENCODING CARBOXYPEPTIDASE OF ASPERGILLUS NIGER
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESS: No. 57053760 No. 5705376disk of No. 5705376th America, Inc.
; STREET: 405 Lexington Avenue, Suite 6400
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/608,224
; FILING DATE: 28-FEB-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/309,341
; FILING DATE: 20-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowney, Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 4247.000-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 867 0123
; TELEFAX: 212 867 0298
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2068 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Aspergillus niger
; FEATURE:
; NAME/KEY: intron
; LOCATION: 572..632
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join (571..633)
; US-08-608-224-1

```

Query Match	4.6%;	Score 76;	DB 1;	Length 2068;
Best Local Similarity	55.8%;	Pred. No. 8.9e-13;		
Matches 145;	Conservative 0;	Mismatches 115;	Indels 0;	Gaps 0;
253	TGGTTCTTCGAAAGCCAGACATAAACCCAGAAAACTGCACCTATCATCTGTGGTTGAAATGGT	312		
633	TGGTTCTTCGAGTCTCGCAATGACCCCGAAGATGATCCCGTTGTTCTGTGGCTCAACCGGT	692		
313	GGCCCTTGGAAAGGATTCTTTGATTCGGTCTCTCGAAGAGTTGGGCCCTTGCCTATGTCAT	372		
693	GGCCCTTGGGTGTCTTTCCCTCACCGGTCTCTTCATGGAGCTTGCGCCCTACAGCATCAC	752		
373	TGGACTTTTGTGACTATCATCAATCAACCCCTCACTCGTGGAAACGAGGTCTCCAAATTACTATT	432		
753	AAGAAGATCCAGCCGGTCTACAATGACTACGCTTGGAACTCCAAACGGCTCCGTGATCTTC	812		
433	CTGTCCAGCCOATGGGAGTCCGGCTTTTCATATAGTGCATACGGTTGATGGTCCATTAAAC	492		
813	CTTGACCAACCTGTCTCAATGTCCGTTTACTCTCTCAGTAACTCTGCTGCAGGACACCGGTC	872		
493	CCTGTAACTGGGGTCTGTCGA	512		
873	GCTGTGSCAAGGACGCTTA	892		

RESULT 13

US-08-967-149-1
: Sequence 1, Application US/08967149
: Patent No. 5939305
: GENERAL INFORMATION:
: APPLICANT: Yaver, Debbie Sue
: APPLICANT: Thompson, Sheryl Ann
: TITLE OF INVENTION: GENE ENCODING CARBOXYPEPTIDASE OF
: TITLE OF INVENTION: ASPERGILLUS NIGER
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: No. 5939305o No. 5939305disk of No. 5939305th America, Inc.
: STREET: 405 Lexington Avenue, Suite 6400
: CITY: New York
: STATE: New York
: COUNTRY: U.S.A.
: ZIP: 10174-8401
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
: CURRENT APPLICATION DATA: US/08/967,149
: APPLICATION NUMBER: US/08/967,149
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/608,452
: FILING DATE: 28-FEB-1996
: APPLICATION NUMBER: US 08/309,341
: FILING DATE: 20-SEP-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Lowney, Karen A.
: REGISTRATION NUMBER: 31,274
: REFERENCE/DOCKET NUMBER: 4247.000-US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 212 867 0123
: TELEFAX: 212 867 0298
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2068 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: Genomic DNA
: ORIGINAL SOURCE:
: ORGANISM: Aspergillus niger
: FEATURE:


```
; NAME/KEY: intron
; LOCATION: 572..632
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join (571..633)
US-08-967-149-1

Query Match
Best Local Similarity 4.6%; Score 76; DB 2; Length 2068;
Matches 145; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

QY 253 TGGTCTTCGAGCCAGACATTAACCCAGAAATGCAACCTATCACATTTGGTTGAATGTT 312
Db |||||
QY 633 TGGTCTTCGAGCTCGCAATGACCCGAGAAATGATCCGTTGTTCTGTGGCTGAAGGT 692
Db |||||
QY 313 GGCCTCGGAAGGATTTCTTGTATCGTCTCTTCGAAGAGTTGGGCCCTTGCCATGTCAT 372
Db |||||
QY 693 GGCCTGGGTGCTCTTCCCTACCGGTCTCTTCATGGAGCTTGGCCCTAGCAGCATCAAC 752
Db |||||
QY 373 TCGACTTTTGATGACTACATCAACCTCTCACTCGTGGAAACGAGGTCTCCAAATTTACTATTC 432
Db |||||
QY 753 AAGAAGATCCAGCGGTCTACAAATGACTACGCTTGGAACTCCAAACGCTCGGTGATCTTC 812
Db |||||
QY 433 CTGTCCCGAGCCATTTGGAGTGGCTTTTTCATATAGTATGATCGTTGATGGGTGCCATTAAC 492
Db |||||
QY 813 CTTGACCAAGCCCTGTCATATGTCGTTTACTCTCTACAGTAACTCTGCTGTGTCAGCGACACGTC 872
Db |||||
QY 493 CTTGTAACCTGGGTGCTGCGA 512
Db |||||
QY 873 GCTGCTGCAAGGACGTCTA 892
Db |||||
```

```
RESULT 14
US-08-899-324-32
; Sequence 32, Application US/08899324
; Patent No. 5945329
; GENERAL INFORMATION:
; APPLICANT: Breddam, Klaus
; APPLICANT: Keilland-Brandt, Morten
; APPLICANT: Mortensen, Uffe
; APPLICANT: Olesen, Kjeld
; APPLICANT: Stenmicke, Henning
; APPLICANT: Wagner, Fred
; TITLE OF INVENTION: CUSTOMIZED PROTEASES
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt
; STREET: 3100 No. 5945329west Center, 90 S. 7th Street
; CITY: Minneapolis
; STATE: MN
; COUNTRY: U.S.A.
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/899,324
; FILING DATE: 23-JUL-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/329,892
; FILING DATE: 27-OCT-1994
; APPLICATION NUMBER: 08/144,704
; FILING DATE: 28-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Kettleberger, Denise M
; REGISTRATION NUMBER: 33,924
; REFERENCE/DOCKET NUMBER: 8648.44USCI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612/332-5300
; TELEFAX: 612/332-9081
```

```
TELEX:
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2632 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 696...2291
; OTHER INFORMATION:
US-08-899-324-32

Query Match
Best Local Similarity 4.0%; Score 66.2; DB 2; Length 2632;
Matches 140; Conservative 0; Mismatches 123; Indels 0; Gaps 0;

QY 241 CATACCTTCTTCTGCTTCTTCGAAAGCCAGACATTAACCCAGAAATGCAACCTATCACATTTG 300
Db |||||
QY 1113 CATTCCTTCTTGGACTTTTGAAGTAGAAACGATCTCTGCAAGGATCCGGTCACTCTT 1172
Db |||||
QY 301 TGGTTGAATGTTGGCCCTCGGAAGCGATTTCTTTGATGGTCTCTTCGAAGAGTTGGGCCCT 360
Db |||||
QY 1173 TGGTTGAACGGGGTCCAGGTGTCTTCTCACTAACCGGGGTGTTCTTTTGAATTAGGACCC 1232
Db |||||
QY 361 TGCCATGTCAATTCGACTTTTGTATGACTACATCAACCTCACTCTGCGAAAGAGTCTGCC 420
Db |||||
QY 1233 TCATCCATTGGACCTGATTTTGAACCCCATCGGAACCCCTTACTCTTGAACAGCAATGCC 1292
Db |||||
QY 421 AATTACTATTCTCTGTCGCCAGCCATTTGGGAGTCCGGCTTTTCATATAGTATGATACGGTTGAT 480
Db |||||
QY 1293 ACCGTGATCTCTCTTGACCGCTGTCAACGTTGGTTCTCGTATTCGGGTCTCTCAGGT 1352
Db |||||
QY 481 GGGTCCATTAACCTCTGTAACCTGG 503
Db |||||
QY 1353 GTTCCCAACACTGTGCGCGCTGG 1375
Db |||||

RESULT 15
US-08-829-892B-32
; Sequence 32, Application US/08329892B
; Patent No. 6187579
; GENERAL INFORMATION:
; APPLICANT: Breddam, Klaus
; APPLICANT: Keilland-Brandt, Morten
; APPLICANT: Mortensen, Uffe
; APPLICANT: Olesen, Kjeld
; APPLICANT: Stenmicke, Henning
; APPLICANT: Wagner, Fred
; TITLE OF INVENTION: CUSTOMIZED PROTEASE
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt
; STREET: 3100 No. 6187579west Center, 90 S. 7th Street
; CITY: Minneapolis
; STATE: MN
; COUNTRY: U.S.A.
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/329,892B
; FILING DATE: 27-OCT-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
```

APPLICATION NUMBER: 08/144,704
FILING DATE: 28-OCT-1993
ATTORNEY/AGENT INFORMATION:
NAME: Kettleberger, Denise M
REGISTRATION NUMBER:
REFERENCE/DOCKET NUMBER: 8648.44US01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612/332-5300
TELEFAX: 612/332-9081
TELEX:

INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 2632 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE:
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 696...2291
OTHER INFORMATION:

US-08-329-892B-32

Query Match 4.0%; Score 66.2; DB 3; Length 2632;
Best Local Similarity 53.2%; Pred. No. 9.4e-10;
Matches 140; Conservative 0; Mismatches 123; Indels 0; Gaps 0;
QY 241 CATACCTTCTTGGTTCCTCGAAGCCAGACATACCCAGAACTGCACCTATCACATTG 300
DB 1113 CATTCTCTTTTGGACTTTTGAAGTAGAAACGATCTCGAAAGGATCCGGTCATCCTT 1172
QY 301 TGGTTGAATGGTGGCCCTCGAAGCGATTCTTTGATCGGTCTCTTCGAAGAGTTGGGCCCT 360
DB 1173 TGGTTGAACGGGGTCCAGGTGTTCTTCACTAACGGGCTGTTCTTTGAATTAGGACCC 1232
QY 361 TGCATGTCAATTCGACTTTTATGATGATCATCAACCCCTCACTCGTGGAAACGAGGTCTCC 420
DB 1233 TCATCCATTGGACCTGATTTGAAACCCATCGGGAACCCCTTACTCTTTGGAACAGCAATGCC 1292
QY 421 AATTACTATTCCTGTCGCCAGCAATGGGAGTCGGCTTTTCATATAGTATACGGTTGAT 480
DB 1293 ACGTGTATCTTCTCGACCAAGCCTGTCAACGTTGGGTCTCTGPAATTCGGGTCCTCAGGT 1352
QY 481 GGGTCCATTAAACCCCTGTAACCTGG 503
DB 1353 GTTCCACACACTGTCGCCGCTGG 1375

Search completed: July 29, 2004, 11:31:23
Job time : 120 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 29, 2004, 11:28:14 ; Search time 791 Seconds
(without alignments)
10339.365 Million cell updates/sec

Title: US-09-712-338-1

Perfect score: 1668

Sequence: 1 atcgctggctacgaattctt.....ccagtggtgtagcatag 1668

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3222919 seqs, 2451570024 residues

Total number of hits satisfying chosen parameters: 6445838

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	374.8	22.5	1859	16	US-10-369-493-36581 Sequence 36581, A
2	91.6	5.5	1686	16	US-10-369-493-27042 Sequence 27042, A
3	70.6	4.2	2218	13	US-09-993-192-1 Sequence 1, Appli
4	70.6	4.2	2218	15	US-10-206-619-1 Sequence 1, Appli
5	67	4.0	1619	16	US-10-369-493-27598 Sequence 27598, A
6	66.2	4.0	1599	16	US-10-369-493-25613 Sequence 25613, A
7	65.6	3.9	1289	16	US-10-369-493-36920 Sequence 36920, A
8	61.8	3.7	1527	16	US-10-369-493-25156 Sequence 25156, A
9	56.4	3.4	1814	9	US-09-736-457-319 Sequence 319, App
10	56.4	3.4	1814	9	US-09-902-941-319 Sequence 319, App
11	56.4	3.4	1814	9	US-09-849-626-319 Sequence 319, App
12	56.4	3.4	1814	10	US-09-476-300-319 Sequence 319, App
13	56.4	3.4	1814	13	US-10-283-017-319 Sequence 319, App
14	56.4	3.4	1814	15	US-10-017-754-319 Sequence 319, App

15	56.4	3.4	1814	15	US-10-113-872-319	Sequence 319, App
16	56.4	3.4	1815	9	US-09-880-107-2296	Sequence 2296, Ap
17	56.4	3.4	1815	17	US-10-450-826-69	Sequence 69, Appl
18	55.8	3.3	710	16	US-10-369-493-29440	Sequence 29440, A
19	54.8	3.3	2864	13	US-09-925-302-331	Sequence 331, App
20	54.8	3.3	2864	13	US-09-925-302-331	Sequence 331, App
21	53.4	3.2	1503	17	US-10-437-963-97753	Sequence 97753, A
22	53.2	3.2	1383	17	US-10-437-963-62737	Sequence 62737, A
23	52.6	3.2	1514	13	US-10-425-114-29826	Sequence 29826, A
24	52.6	3.2	1905	13	US-10-424-599-104947	Sequence 104947, A
25	51.2	3.1	1326	16	US-10-369-493-27091	Sequence 27091, A
26	49	2.9	1113	16	US-10-369-493-36625	Sequence 36625, A
27	49	2.9	1574	13	US-10-425-114-6508	Sequence 6508, Ap
28	49	2.9	1717	13	US-10-424-599-132435	Sequence 132435, A
29	47.6	2.9	390	9	US-09-790-399-7	Sequence 7, Appli
30	47.6	2.9	1551	14	US-10-084-018-4	Sequence 109, App
31	47.6	2.9	1684	9	US-09-729-674-109	Sequence 39, Appl
32	47.6	2.9	1697	10	US-09-796-753-39	Sequence 33, Appl
33	47.6	2.9	1949	10	US-09-971-429B-33	Sequence 586, App
34	47.6	2.9	1963	13	US-10-296-115-586	Sequence 163, App
35	47.6	2.9	2076	9	US-09-909-320-163	Sequence 163, App
36	47.6	2.9	2076	9	US-09-905-088B-163	Sequence 163, App
37	47.6	2.9	2076	9	US-09-905-291A-163	Sequence 163, App
38	47.6	2.9	2076	9	US-09-902-853-163	Sequence 163, App
39	47.6	2.9	2076	9	US-09-907-824-163	Sequence 163, App
40	47.6	2.9	2076	9	US-09-907-841-163	Sequence 163, App
41	47.6	2.9	2076	10	US-09-904-011-163	Sequence 163, App
42	47.6	2.9	2076	10	US-09-906-742-163	Sequence 163, App
43	47.6	2.9	2076	10	US-09-906-838-163	Sequence 163, App
44	47.6	2.9	2076	10	US-09-907-613-163	Sequence 163, App
45	47.6	2.9	2076	10	US-09-907-942-163	Sequence 163, App

ALIGNMENTS

RESULT 1

US-10-369-493-36581
; Sequence 36581, Application US/10369493
; Publication No. US2003023675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 36581
; LENGTH: 1859
; TYPE: DNA
; ORGANISM: Aspergillus nidulans
US-10-369-493-36581

Query Match 22.5%; Score 374.8; DB 16; Length 1859;
Best Local Similarity 58.0%; Pred. No. 3.2e-114;
Matches 1067; Conservative 0; Mismatches 402; Indels 371; Gaps 8;

QY	76	TCGTCGGTAGAAGACAGCTCCCAAGAACCCACCGGGGTCAAGACTCTTACACCGCA	135
Db	19	TCCTTCAACGACGACGACTACCGCGGACCACTGGTGTCAAGAAATTCAGACTCC	78
QY	136	ACAATGTCCATCCGTTACAGGAACCCGGGGCAGAGGGGTCTGCGAGACTACCCG	195
Db	79	AATGGTGTGACCATCGATACCAAGAGGCTGGTAAAGAGGGTGTCTGTGAGAGCCCT	138
QY	196	GGTGTCAATCTACTCTCGATATGCGACACCTCTCCGAGTCCCATACCTTCTTCTGG	255

Db 139 GGTGTCTAACTCTTATCTCGGATACGTTGATCTATCGCTGAGTCTCACACCTTCTTTTAC 198
QY 256 TTCTTGAAGCAGACATAACCCAGAACTGCACCTATCACATGTGTGTTGAATGGTGGC 315
Db 199 TTCTTGAAGCTGCTCATGACCCGCAACGCTCTATCACTCTGTGGCTGAATGGTGGC 258
QY 316 CTTGGAAGCGATTTCTTCTGATCGGTCTCTTGA 347
Db 259 CCGGAAGTGATTCGCTTATCGGGCTTTTGAAGGTTTGTACTGCTGACTGGGTTACAA 318
QY 348 -----AGATTGGGCCCTTGGCCATGTCATGTCGACTT 379
Db 319 GAAGTGTGATGACTGATGGGCTTCTTTAGAACTGGGACCGTGCCTATCAGGAGAACC 378
QY 380 TTGATGACTACATCAACCTCTCACTGTTGAACGAGGTCTCCAAATTTACTATTCTCTGTC 439
Db 379 TGACGACCTACGTCAACCATACTCGTGAATGAAGTCTCCAACTTTTGTTCATCTGCG 438
QY 440 AGCCATTTGGAGTC-----ATA 453
Db 439 AGCCATTTGGAGTCGATGAGGAATTTGTCTTTTGTGTCGTCAGCTTCTCACTGG 498
QY 454 -----GGCTTTTCATATAGTATACGTTGATGGTCCATTAAACCTGTAACTGGGTC 507
Db 499 GCCATGAGATTTCTGATGCGGAACAGAGGAAGGTCTGCTCAATCCCGTCAACCGGTG 558
QY 508 GTCGAAATTCGAGCTTTGACGAGTTTCAGGCGCGGTACCCAACTGATGATGCCACTCTG 567
Db 559 GTGAGGATAGTTCTTCTGCTGCGTCCAGGTCGCTATCCCAAAATCAACGCTACCCCTG 618
QY 568 ATCG-----ATA 574
Db 619 ACCGGTAATTTCTCAACAACTGTCATTTCTCTAGTTTAGAGCTAAAGCTGACTAGATA 678
QY 575 CTACCAATCTTCCGACAGGCGCTTGGGAGATCTGCAAGATTCTCTAGTGGACTAC 634
Db 679 CCACCGAGTCCCGGGAAGGCACCTTGGGAAGTCTGCAAGGCTTCTCGGTGGTCTAC 738
QY 635 CTAGCTTGACTCTAGGTCAGTCTTAAGGACTTCAAGTCTATGGACGAGAGCTATGGAG 694
Db 739 CTCAGTTGACAGCACTATCAAGTCCAAAGACTTTAATTTATGAGCTGAGAGTTACGGAG 798
QY 695 GGCATATGGTCTGATCTTCAATCAATTTTACGACGAGATGAGGAATGCCAAGC 754
Db 799 GACATACGCTCCCGAATCTTCAATCACTTTTACGACGAGAACCAAAATTCGCAAGC 858
QY 755 GTAGTGTAAATGGTTCAGCTTAATTTCAACTCTCTGGGAATTTAAGCGCATCATCG 814
Db 859 GAACCATTTGATGGTATCGAGCTCACTTCAACTCACTGGGAATTTAATGTTATCATCG 918
QY 815 ACAGGCGGAT-----C 825
Db 919 ATGAGGGGATTCAGGTACGGATAGACAAGCCGCTCAGCAATATCTCTGACTAGAAAC 978
QY 826 CAGGCGCTTACTACCTGATTCGTTGCAACAACTACCTACGTTACAGGC-----878
Db 979 TAGGCACCTTTCTACCTGATTTGGCGTGCAACAACCTACGGCATCAAGCGGTGAGT 1038
QY 879 -----TGTCACGAGACCG 892
Db 1039 ACTGTCTTTTACGAATGACGACATACTTACAAGCTGAAACAGGTTTACGAGACAG 1098
QY 893 TCTACAACTCATGAGTTGGCAACCAATGCAATGGTTGCGAGGATTTGATTTCCA 952
Db 1099 TCTACAACTCATGAAATTTGGCAACGAGATGGGCAACGATGCCAGGATTCAGATGGCCC 1158
QY 953 CTTGCAACAGACAAACCGCACCGCATTTAGCTAGCTAGCGCTCTGCGCGCAAGCCACCA 1012
Db 1159 TCTGCAAGCAACAAACCGTTCTTTTGGCGGACTATGCTCTCTGCTCGGAGGCACTA 1218
QY 1013 ACATGTGACGGGCAATGTTGAGGGGCGCATCTAC-----1047

Db 1219 ACATGTGACGAGATAACGTTGGTACGTACTTCTACTCTGAGACTCTGGGAAAGTAATCAGC 1278
QY 1048 -----GCCTTTGCTGGTCTGCTGTGTANGA 1073
Db 1279 TGATTATCTTCCACAGAGGACTCTCTACTACAGGCTACGAGCAACTCGAGGCGTATATGA 1338
QY 1074 TATTCCGCATCCATATGAT-----1092
Db 1339 CATTGCGCACCGGCACATGTAAAGCCGTTGATCTTAATCCAGGTTCTCTACTATAACATA 1398
QY 1093 -----GACCCGACTCCGCCAAGTTATTAACAAATTTCTGGCAAGGACTCTGTGTAT 1145
Db 1399 GGAAGGAGACCAACACACCCCGAACCTCTTCGTCCTCTCCAGCAAGACTTGGGTGAT 1458
QY 1146 GGAAGC-TATGGCGTCAAGATCAACTACACCGAGTCCCAATTAAGACCTCTACTACGCTT 1204
Db 1459 GAATGCTTCTTGGCGTGAACATCACTACCTCTTCCAAATAACGAAATTTACTATGCTT 1518
QY 1205 TCCAGCAACAGGCGACTTTGTCTGGCCCAACTTTCATCGAAGACCTCGAGGAGATCTCTTG 1264
Db 1519 TCCAGCAACCGGTGATTTCTGTTGGCCCAACTTTCATCGAGGATCTCGAGGAAATCTCTCA 1578
QY 1265 CTCTCCCGTGGTGTCTCTCTATGCGGAGCGCGATTAATCTGCAACTGTTTCG 1324
Db 1579 CTTACCCAGTCCGCTCAGTCTGATCTATGGCGACGAGACTACATCTGCAACTGTTTCG 1638
QY 1325 GGGTCAGCGCTTTCCCTCGCTCGAACTCTCCCAAGCCGCCAGTTCCGAAGCGCAG 1384
Db 1639 GAGGCCAAGCGCTCTCCCTCGCCGTCAGAGTACCCGATTCAGGAGTTCCAGGCGCGCG 1698
QY 1385 GGTACACGCCCCCTGAAAGTCAACGCGCTCGAGTATGGGGAACCTCGCGAGTATGGTAAT 1444
Db 1699 GATACACGCTCTTCTTGTGATGCGTGGAGTACGAGAGACTCGAGAGTATGGCAAT 1758
QY 1445 TCTCTTCACTCGCTCTATGAGGCGAGCGCATGAAGTCCCATCTACTACAGCCCATCGCT 1504
Db 1759 TCTCTTCTCCGCTGTCTATGAGGCGCGCCAGGTTCTTACTATCAGCCTGCTGCT 1818
QY 1505 CCCTGCAATTTGTTAAACGAGCTATCTTCGTTGGGATAT 1544
Db 1819 CGCTGAGCTTTTAAACGCACTTTGTTGGTGGGATAT 1858

RESULT 2

US-10-369-493-27042
; Sequence 27042, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 27042
; LENGTH: 1686
; TYPE: DNA
; ORGANISM: Neurospora crassa
US-10-369-493-27042

Query Match 5.5%; Score 91.6; DB 16; Length 1686;
Best Local Similarity 49.4%; Pred. No. 4.2e-19;
Matches 312; Conservative 0; Mismatches 299; Indels 21; Gaps 2;
QY 749 CCAACGGTAGTGTAAATGGTGTTCAGCTTAATTTCAACTCTCTGGGAATTAATAACGGCA 808

Db 842 CCAAGGCTGGCAGTAGTAACCGGAAATATAAACTCACCTTGGCATCATGAACGGCT 901
QY 809 TCATCGAGGAGGATCCAGGCGCTTACTACCTGATTCCTGTGAACATACCTACG 868
Db 902 GCATCGAGGCTGCTCAAGGCTCCCGTACGTCGAGATGCCATCAACACACGTCAG 961
QY 869 GTATCAAGGCTGTCAACGAGACCGTCTACAACTACATCAAGTTTGGCCAAACCAATGCC-- 926
Db 962 GCATCAAGCTGATCGACCAAGCCAGCGCGGAGCCATCCTCGGGCTTCAACGCCCTG 1021
QY 927 -AATGGTTGGCAGATTTTCCACTGCAACAGACAAACCGCACCCATTAAGTCG 985
Db 1022 ACAGGGGCTGCAAGAGACCTGATCTCGCTTCCCGCAAGCCCAAGCCGCTCTCGACCGC 1081
QY 986 ACTAGCCCTCTGGCGGAGCCACCAACATGTGACGGGAATGTTGAGGGGCACTACT 1045
Db 1082 TGGACCAAGGTACCGACGAAACCGTCAACCAAGGCTCGGCCCTACGCTCGCTACCTGCC 1141
QY 1046 ACGCCTTTGC-----TGGTCGTGGTGTCTATGATATTCGGCATCCAT 1087
Db 1142 AGCAGCTCTCGGCTCGCTCGCCAGTGGCGTAAATGCCCTACGACATTCGGCACATGG 1201
QY 1088 ATGATGACCGGACTCCGCGCAAGTTATTACAACTTTTGGCAAGAGCTCTCTCATGG 1147
Db 1202 GTCCCGATGCGTTTCCCGGATTAACCACTACCTCGAGTATCTCAACTCGCGCGCGTGAAG 1261
QY 1148 ACGTATCGGCGTCAACATCAACTACACCCAGTCCATATGACGTCTACTACGCTTCC 1207
Db 1262 AGCGATCGGCTCGTGGTCAACTACACGGATATTCGCGCGTAGTCTACCAAGCCTTTT 1321
QY 1208 AGCAACAGCGGACTTTGTCTGGCCCACTTCATCGAAGACCTCGAGGAGATCCTTGCTC 1267
Db 1322 TCCAAACCGGGATAGGCTCGCGCGGTTGATCGCCAGCTAGCTTCGTTCTGCAGC 1381
QY 1268 TCCCGTGGTGTCTCCCTCATCTATGGCGAGCGCGATTAATCATGCAACTGGTTCCGGC 1327
Db 1382 CGGCGTGGCATCGGCTCTGGTGTATGGCGACCGGATTAATCATGCAACTGGATGGCG 1441
QY 1328 GTACAGGCGGTTTCCCTCGCTCGGAACCTACTCC 1359
Db 1442 GCGAAGCGGTCTCTGGCTCTTCCGACGCC 1473

RESULT 3

US-09-993-192-1
; Sequence 1, Application US/09993192
; Publication No. US20020150983A1
; GENERAL INFORMATION:
; APPLICANT: Korea Research Institute of Bioscience and Biotechnology
; APPLICANT: Dong Kook Pharmaceutical Co.
; APPLICANT: Rhee, Sangki
; APPLICANT: Choi, Euisung
; APPLICANT: Kang, Hyunah
; APPLICANT: Sohn, Junghoon
; APPLICANT: Bae, Junghoon
; APPLICANT: Kim, Moowoong
; APPLICANT: Agaphonov, Michasel
; TITLE OF INVENTION: Hansenula polymorpha mutants and process for the preparation of
; TITLE OF INVENTION: recombinant proteins using the same
; FILE REFERENCE: 4220-116 US
; CURRENT APPLICATION NUMBER: US/09/993,192
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: US/09/674,617A
; PRIOR FILING DATE: 2001-01-03
; PRIOR APPLICATION NUMBER: PCT/KR00/00173
; PRIOR FILING DATE: 2000-03-04
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 2218
; TYPE: DNA
; ORGANISM: Hansenula polymorpha DLI
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: US/09/993,192
; CURRENT FILING DATE: 2001-11-14
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: US/09/674,617A
; PRIOR FILING DATE: 2001-01-03
; PRIOR APPLICATION NUMBER: PCT/KR00/00173
; PRIOR FILING DATE: 2000-03-04
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 2218
; TYPE: DNA
; ORGANISM: Hansenula polymorpha DLI
; PUBLICATION INFORMATION:

; DATABASE ACCESSION NUMBER: U67174
; DATABASE ENTRY DATE: 1996-08-17
; RELEVANT RESIDUES: (1)..(2218)
US-09-993-192-1

Query Match 4.2%; Score 70.6; DB 13; Length 2218;
Best Local Similarity 53.4%; Pred. No. 5.6e-12;
Matches 148; Conservative 0; Mismatches 129; Indels 0; Gaps 0;
QY 241 CATACCTTCTTCTGGTTCTTCGAAGCCAGACATCAACCCAGAAACTGCACCTATCACATTG 300
Db 840 CATTTCTTCTATGGATGTTGAGTCCAGAAATGACCCGGTCAACGACCTGTGATTCTG 899
QY 301 TGGTTGAATGGTGGCCCTCGAAGCGATCTTTTGAATCGGTCTCTTCGAAGAGTTGGGCCCT 360
Db 900 TGGCTCAACGGTGGTCCAGGATGCTCTTCCTTGACTGGAATGCTTTTGGAGCTCGGCTCT 959
QY 361 TGCCATGTCGAATTCGACTTTTGATGACTACATCAACCTCACTCGTGAAGAGGTCCTCC 420
Db 960 GCTTCTATCGGTCCAGATCTCAAGCCAAATCAACCAACCCATATTCGTGGAATTCCAATGCC 1019
QY 421 AATTTACTATTCTCTGCCAGCCATTCGGAGTCCGGCTTTTCATATAGTATACGGTTGAT 480
Db 1020 ACTGTGATTTTCTTGACCGCTGTCATGTTGGATTCTCGTACTCTTCCAAGTCTGTT 1079
QY 481 GGGTCCATTAACCTCTGTAATCTGGGTCGTGCAAAAT 517
Db 1080 TCTAACACGTCGCGAGCTGGTAAAGACGTCTATGCTT 1116

RESULT 4

US-10-206-619-1
; Sequence 1, Application US/10206619
; Publication No. US2003010898A1
; GENERAL INFORMATION:
; APPLICANT: Korea Research Institute of Bioscience and Biotechnology
; APPLICANT: Dong Kook Pharmaceutical Co.
; APPLICANT: Rhee, Sangki
; APPLICANT: Choi, Euisung
; APPLICANT: Kang, Hyunah
; APPLICANT: Sohn, Junghoon
; APPLICANT: Bae, Junghoon
; APPLICANT: Kim, Moowoong
; APPLICANT: Agaphonov, Michasel
; TITLE OF INVENTION: Hansenula polymorpha mutants and process for the preparation of
; TITLE OF INVENTION: recombinant proteins using the same
; FILE REFERENCE: 4220-116 US
; CURRENT APPLICATION NUMBER: US/10/206,619
; PRIOR FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: US/09/674,617
; PRIOR FILING DATE: 2001-01-03
; PRIOR APPLICATION NUMBER: PCT/KR00/00173
; PRIOR FILING DATE: 2000-03-04
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 2218
; TYPE: DNA
; ORGANISM: Hansenula polymorpha DLI
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: U67174
; DATABASE ENTRY DATE: 1996-08-17
; RELEVANT RESIDUES: (1)..(2218)
US-10-206-619-1

Query Match 4.2%; Score 70.6; DB 15; Length 2218;
Best Local Similarity 53.4%; Pred. No. 5.6e-12;
Matches 148; Conservative 0; Mismatches 129; Indels 0; Gaps 0;
QY 241 CATACCTTCTTCTGGTTCTTCGAAGCCAGACATCAACCCAGAAACTGCACCTATCACATTG 300
Db 840 CATTTCTTCTATGGATGTTGAGTCCAGAAATGACCCGGTCAACGACCTGTGATTCTG 899

QY 301 TGGTTGATGTCGGCTGGAGCGATCTTTTGATCGTCTCTTCGAGAGTGGGCGCT 360
 Db 900 TGGCTCAACGGTGGCCAGGATCTCTTCCTTGATCTGGAATGCTTTTGACCTGGCTCT 959
 QY 361 TGCCATGTCAATTCGACTTTTGATGACTACATCAACCTCACTCGTGGAAAGAGTCTCC 420
 Db 960 GCTTCTATCGGTCCAGATCTCAAGCAATCAACCACTATCTGTGGAATTCGAATGCC 1019
 QY 421 AATTACTATTCCTGTCAGCAATGGAGTCGGCTTTTCATATAGTATACGTTGAT 480
 Db 1020 ACTGTGATTTCTTCAGCAGCTGTCAATGTTGGAATCTCGTACTCTTCCAAAGTCTGT 1079
 QY 481 GGGTCCATTAACCTGTAACCTGGGCTGTCGGAAT 517
 Db 1080 TCTAACACGGTCGAGCTGGTAAGACGCTCTATGCTT 1116

RESULT 5
 US-10-369-493-27598
 ; Sequence 27598, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; CURRENT FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; PRIOR FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 27598
 ; LENGTH: 1619
 ; TYPE: DNA
 ; ORGANISM: Neurospora crassa
 US-10-369-493-27598

Query Match 4.0%; Score 67; DB 16; Length 1619;
 Best Local Similarity 54.7%; Pred. No. 7.5e-11;
 Matches 133; Conservative 0; Mismatches 110; Indels 0; Gaps 0;
 QY 254 GGTTCCTCGAAGCCAGACATAACCCAGAACTGCACCTATCACTTGTGGTGAATGGTG 313
 Db 466 GGTTCCTCGAGTCTCGCAACGATCCCAAGAACGACCCCGCTGTTTGTGGTTGAACGGTG 525
 QY 314 GCCTCGAAGCGATCTTTGATCGGTCTCTTCGAGAGTGGGCCCTTCCCATGTCAATT 373
 Db 526 GCCCGCGCTGCTCTTCGCTACTGTGCTCTCTTCGAGCTCGGCCCTTCTTCATGACA 585
 QY 374 CGACTTTTGTGACTACATCAACCCCTCACTCGTGGAAAGAGGTCTCCAAATTTACTATCC 433
 Db 586 AAGAGCTCAGGTTGTGACCAAGAGTACGCTTGGAAACAAACGACGAGTATCTTCC 645
 QY 434 TGTCAGCAATGGGAGTCGGCTTTTCATATAGTATACGTTGATGGTTCATTAACC 493
 Db 646 TTGACCAGCCCGTCAACGTCGGGTACTCTGCTACTCCGGCAACGCGTCAGCAACACTGTG 705
 QY 494 CTG 496
 Db 706 CTG 708

RESULT 6
 US-10-369-493-25613
 ; Sequence 25613, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; CURRENT FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; PRIOR FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 25613
 ; LENGTH: 1599
 ; TYPE: DNA
 ; ORGANISM: Saccharomyces cerevisiae
 US-10-369-493-25613

Query Match 4.0%; Score 66.2; DB 16; Length 1599;
 Best Local Similarity 53.2%; Pred. No. 1.4e-10;
 Matches 140; Conservative 0; Mismatches 123; Indels 0; Gaps 0;
 QY 241 CATACCTTCTTCTGGTCTTCGAAGCCAGACATAACCCAGAACTGCACCTATCACATTG 300
 Db 418 CATTTCTTCTTTGGACTTTTGAAGTAGAAGACGATCCTCGAAGGATCCGGTCACTCT 477
 QY 301 TGGTTGAATGTTGGCCCTCGAAGCGATCTTTTGATCGGTCTCTTCGAAAGTGGGCCCT 360
 Db 478 TGGTTGAAGCGGGTCCAGTTGTTCTTCACTAACCGGCTCTCTTTGAATTAGGACCC 537
 QY 361 TGCCATGTCAATTCGACTTTTGTGATGACTACATCAACCTCACTCGTGGAAAGAGTCTCC 420
 Db 538 TCATCCATTGGACCTGATTTGAAACCCATCGGGAACCCCTTACTCTTGGAAACAGCAATGCC 597
 QY 421 AATTACTATTCCTGTCAGCAATGGAGTCGGCTTTTCATATAGTATACGTTGAT 480
 Db 598 ACCGTGATCTTCTTGACCGCTCTCAACGTTGGTCTCTGTAITCCGGTCTCAGGT 657
 QY 481 GGGTCCATTAAACCTGTAACTGG 503
 Db 658 GTTCCACACATGTCGCCGCTGG 680

RESULT 7
 US-10-369-493-36920
 ; Sequence 36920, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; CURRENT FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; PRIOR FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 36920
 ; LENGTH: 1289
 ; TYPE: DNA
 ; ORGANISM: Aspergillus nidulans
 US-10-369-493-36920

Query Match 3.9%; Score 65.6; DB 16; Length 1289;
 Best Local Similarity 53.5%; Pred. No. 1.9e-10;
 Matches 137; Conservative 0; Mismatches 119; Indels 0; Gaps 0;
 QY 254 GGTTCCTCGAAGCCAGACATAACCCAGAACTGCACCTATCACTTGTGTTCAATGGTG 313
 Db 127 GGTTCCTTGAGTCACGAAGAGCCCAAAACGATCCTGTGTTCTGTGTTGACCGTG 186


```
; NUMBER OF SEQ ID NOS: 2157
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 319
; LENGTH: 1814
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-283-017-319

Query Match      3.4%; Score 56.4; DB 13; Length 1814;
Best Local Similarity 52.5%; Pred. No. 2.9e-07;
Matches 148; Conservative 0; Mismatches 131; Indels 3; Gaps 1;

QY 193 CCGGTGTCAAACTCTACTCTGGATATGTCGACACTCTCTCCGAGTCCCATACCTTCTTC 252
Db      |||||
QY 139 CCGTCTTTCCGCGAGTACTCCGGCTACCTCAAAAGCTCCGGCTCCAAGCACTCCACTAC 198
Db      |||||
QY 253 TGGTCTTCGAGCCAGACATACCCAGAACTGCACCTATCACATTTGGTGTGAATGGT 312
Db      |||||
QY 199 TGGTTTGGAGTCCAGAGGATCCGAGAACAGCCCTGTGGTCTTTGGCTCAATGGG 258
Db      |||||
QY 313 GGCCCTGGAAGCGAFTCTTTGATCGGTCTCTTCGAAAGAGTTGGGCCCTTGCCATGCAAT 372
Db      |||||
QY 259 GGTCCCGGCTGCAGCTCACTAGATGGGCTCTCCAGAGCATGGCCCTTCTCTGGTCCAG 318
Db      |||||
QY 373 TCGACTTTTGATGACT---ACATCAACCTCACTCGTGGAAAGAGTCTCCAAATTTACTA 429
Db      |||||
QY 319 CCAGATGGTGTACCCCTGGAGTACAAACCCCTATTCTTGAATCTGATTGCCAATGTGTTA 378
Db      |||||
QY 430 TTCCTGTCCCGAGCATTCGGAGTCCGGCTTTTCATATAGTAGT 471
Db      |||||
QY 379 TACCTGGAGTCCCGAGCTGGGGTGGGCTTCTCTACTCCGAT 420
Db      |||||
```

RESULT 14

```
US-10-017-754-319
; Sequence 319, Application US/10017754
; Publication No. US20030054363A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.478C18
; CURRENT APPLICATION NUMBER: US/10/017,754
; CURRENT FILING DATE: 2001-10-29
; NUMBER OF SEQ ID NOS: 2004
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 319
; LENGTH: 1814
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-017-754-319

Query Match      3.4%; Score 56.4; DB 15; Length 1814;
Best Local Similarity 52.5%; Pred. No. 2.9e-07;
Matches 148; Conservative 0; Mismatches 131; Indels 3; Gaps 1;

QY 193 CCGGTGTCAAACTCTACTCTGGATATGTCGACACTCTCCCGAGTCCCATACCTTCTTC 252
Db      |||||
QY 139 CCGTCTTTCCGCGAGTACTCCGGCTACCTCAAAAGCTCCGGCTCCAAGCACTCCACTAC 198
Db      |||||
QY 253 TGGTCTTCGAGCCAGACATACCCAGAACTGCACCTATCACATTTGGTGTGAATGGT 312
Db      |||||
QY 199 TGGTTTGGAGTCCCGAGAGGATCCGAGAACAGCCCTGTGGTCTTTGGCTCAATGGG 258
Db      |||||
```

```
QY 313 GGCCCTGGAAGCGAFTCTTTGATCGGTCTCTTCGAAAGAGTTGGGCCCTTGCCATGCAAT 372
Db      |||||
QY 259 GGTCCCGGCTGCAGCTCACTAGATGGGCTCTCCAGAGCATGGCCCTTCTCTGGTCCAG 318
Db      |||||
QY 373 TCGACTTTTGATGACT---ACATCAACCTCACTCGTGGAAAGAGTCTCCAAATTTACTA 429
Db      |||||
QY 319 CCAGATGGTGTACCCCTGGAGTACAAACCCCTATTCTTGAATCTGATTGCCAATGTGTTA 378
Db      |||||
QY 430 TTCCTGTCCCGAGCATTCGGAGTCCGGCTTTTCATATAGTAGT 471
Db      |||||
QY 379 TACCTGGAGTCCCGAGCTGGGGTGGGCTTCTCTACTCCGAT 420
Db      |||||

RESULT 15
US-10-113-872-319
; Sequence 319, Application US/10113872
; Publication No. US20030170255A1
; GENERAL INFORMATION:
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.478C19
; CURRENT APPLICATION NUMBER: US/10/113,872
; CURRENT FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 2011
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 319
; LENGTH: 1814
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-113-872-319
```

```
Query Match      3.4%; Score 56.4; DB 15; Length 1814;
Best Local Similarity 52.5%; Pred. No. 2.9e-07;
Matches 148; Conservative 0; Mismatches 131; Indels 3; Gaps 1;

QY 193 CCGGTGTCAAACTCTACTCTGGATATGTCGACACTCTCCCGAGTCCCATACCTTCTTC 252
Db      |||||
QY 139 CCGTCTTTCCGCGAGTACTCCGGCTACCTCAAAAGCTCCGGCTCCAAGCACTCCACTAC 198
Db      |||||
QY 253 TGGTCTTCGAGCCAGACATACCCAGAACTGCACCTATCACATTTGGTGTGAATGGT 312
Db      |||||
QY 199 TGGTTTGGAGTCCCGAGAGGATCCGAGAACAGCCCTGTGGTCTTTGGCTCAATGGG 258
Db      |||||
QY 313 GGCCCTGGAAGCGAFTCTTTGATCGGTCTCTTCGAAAGAGTTGGGCCCTTGCCATGCAAT 372
Db      |||||
QY 259 GGTCCCGGCTGCAGCTCACTAGATGGGCTCTCCAGAGCATGGCCCTTCTCTGGTCCAG 318
Db      |||||
QY 373 TCGACTTTTGATGACT---ACATCAACCTCACTCGTGGAAAGAGTCTCCAAATTTACTA 429
Db      |||||
QY 319 CCAGATGGTGTACCCCTGGAGTACAAACCCCTATTCTTGAATCTGATTGCCAATGTGTTA 378
Db      |||||
QY 430 TTCCTGTCCCGAGCATTCGGAGTCCGGCTTTTCATATAGTAGT 471
Db      |||||
QY 379 TACCTGGAGTCCCGAGCTGGGGTGGGCTTCTCTACTCCGAT 420
Db      |||||
```

Search completed: July 29, 2004, 13:35:46
Job time : 794 secs

Db 67 FFEARHDPANAPITLWMLNGPGSDSLGLFEGLYCLTGLQEVVDVLMFGLKELGPCALSEN 126
QY 127 FDDYINPHSNVEVSNLFLSOPLGVG-----FSYSDTVGDSINPVTGVVENSFF 175
Db 127 LTTVYNPHSNVEVSNLFLSOPLGVGKMKELCLVNSCSFSLGHRIFVCGN--RGRVAQSRH 184
QY 176 AGVQG-----RYPTID-----ATLIDTNNLAAEAWEILOQFELS 209
Db 185 RCRGXFRRWRPGLSQRVYDRXFLQORALSXFRAKAVDITTELAAKATWVVOGFLG 244
QY 210 GLPSLDRVQKDFSLWTESVGGHYGPAFNHFFVEQNERIANGSVNGVQVNFNSLIGLNG 269
Db 245 GLPQLDSTIKSKDFNLWTESVGGHYGPAFNHFFVEQNERIANGSVNGVQVNFNSLIGLNG 304
QY 270 IIDRAIQ-----APYYP-----EFVAVNNTY-----GIXAVN 295
Db 305 IIDREGIQRDRAAQYLSDXKLFTFLPXVCGAQLRHSREYCLFSKWTXTLTSXKQVN 364
QY 296 ETVYNMFKFANOMPNGCODLISTCKQTNRTALADYALCAEATNMCRDNVEGYY----- 349
Db 365 ETVYNMFKFANEMNGCQDIALCKATNRSSFADYALCSEATNMCRDNV--GTIFYLRLWE 423
QY 350 ----APAGRGV-----YDIRHPYD-----DTPPSYXNKFLAK 378
Db 424 SNQLIIFHRGPTTGYEQLEAYMDIRHPHNVSRXSRSFSLTXEKDTPPENLFPVFLQ 483
QY 379 DSVMDAIGVNTYOSNDVYVAFQOTGDFVWPNFIEDLEILALPVRVSLIYGDADYIC 438
Db 484 DWMNASVNVNTSSNEIYVAFQOTGDFVWPNFIEDLEILALPVRVSLIYGDADYIC 543
QY 439 NWFQGOAVSLAANYSOAQRSGAGYTPLKVNVEYGETREYGNFSTRVYRAGHEVPYYQ 498
Db 544 NWFQGOAVSLAVKYPHSGKQFAAGYTPLLVGVEYGETREYGNFSTRVYRAGHEVPYYQ 603
QY 499 PIASLOLNFRTIFG 512
Db 604 PAASLOLNFRTIFG 617
RESULT 2
US-10-369-493-3355
; Sequence 3355, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; NUMBER OF SEQ ID NOS: 47374
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Neurospora crassa
US-10-369-493-3355
Query Match 27.3%; Score 812.5; DB 15; Length 562;
Best Local Similarity 36.0%; Pred. No. 1.9e-67;
Matches 198; Conservative 83; Mismatches 182; Indels 87; Gaps 19;
QY 33 PKNPTGVKVTUTAN--NVTRYKEPAGEGVCEIT--FGVKSYSGVDTSPS--SHTFEPWF 87
Db 28 PYQAVGLKVTSPSNPNIKISYKEP--KDVCTAFRRQKQYGTGVTVPVGPYTHLFPWFV 85
QY 88 EARNHPETAPITLWMLNGPGSDSLGLFELGPGCHV-----NSTFDDYINPHSNVEVSNLL 143

Db 86 GARE--PTSALTWMLNGPGSSSMFGLFAENGPCQVVEKGASRLETAAREWGMDRANML 143
QY 144 FLSQPLGVGFSYSDTVGDSINPVTGVVE-----NSSFA--GVQGRYPTID 186
Db 144 FVDQPNHVGFSYDTPNGSLDLITGMVSPTMQLPDNLPPSLFLNGTFTSSNNIQ----- 196
QY 187 ATLIDTNNLAAEAWEILOQFGLSPS--LDSRVQSKDFSLWTESVGGHYGPAFNHFFVEQ 245
Db 197 -NTANTTQNDAMAVYHLLQGLSTFTFPYVDPKHSPLGVNLFSTESYGGHYGPFVADTWQKE 255
QY 246 NERIANGSV-----NGVQ-----LNFSLGIINGIIDEALQAPYYPEFA 284
Db 256 NDKLSKRAMPLSRDQDANNIQSRPPKAGSSNREIKLTLGIMNGCDDLLVQGSRYVEMA 315
QY 285 VNNYTGKAVNTEVYNTMFKFANOMPNGCODLISTCKQ-----TNRALADYAL 332
Db 316 INNTYGIKLIDQATADAITSGFNAPDTGCKDLIILACQAQALDPLDQGTDTV---NQV 372
QY 333 CAEATNMCRDNVEGYYAFAGRGVYDIRHPYDDTPPSYXNKFLAKDSVMDAIGVNTY 392
Db 373 CAYASLTQO--QLLSGVLA--SGANAYDIAHMGPDAPPDYHLYEYLSNRVQEAIGSVVNYT 430
QY 393 QSNNDVYVAFQOTGDFVWPNFIEDLEILALPVRVSLIYGDADYICNWFQGOAVSLA--- 449
Db 431 DISPVYQAFQOTGDRARGGLIAKLASLLQVRIGLYGDRDYICNMWGGEAVSLALAD 490
QY 450 -----ANYSOAQFRSGAGYTPLKVNVEY--GETREYGNFSTRVYRAGHEVPYQPIAS 502
Db 491 AMEDLPKSPYTPKPOAGYENIQTNREVGVGVVRFQGNLSFSLYQSGHFVPAYQETA 550
QY 503 LQLFNRTIFG 512
Db 551 PRVFERIIGG 560
RESULT 3
US-10-369-493-1469
; Sequence 1469, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; NUMBER OF SEQ ID NOS: 47374
; LENGTH: 508
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-10-369-493-1469
Query Match 21.3%; Score 636; DB 15; Length 508;
Best Local Similarity 32.6%; Pred. No. 7.6e-51;
Matches 167; Conservative 70; Mismatches 197; Indels 78; Gaps 14;
QY 19 LPQSTPASVGRRLQ----PKNPTGVKVTUTANNTYRKEPAGEGVCEITTPGVKSYSGYV 74
Db 43 LPQNTQOTLKLDRKLNHDDDLFTFTTSSVDTSYLRRTVDPKSLGI---DTVKQWSGYM 98
QY 75 DTSPESHTFWFPEARNPETAITLWMLNGPGSDSLGLFELGPGCHVNSTFDDYINPH 134
Db 99 DYKDSKHFFWFPEESRNDPANDPILWMLNGPGCSFTGLLPELGPSSIGADMKPIHNP 158
QY 135 SWNEVSNLLPLSOPLGVSFSDTVGDSINPVTGVVENSFAGVQGRYPTIDATLIDTNN 194

Db 159 SWNNASMIFFLEQPLGVGFSGD-----EKVSSTK 188
 QY 195 LAEEAAWIILOFLSGLSDSRVQSKDFSLWTSYSGHYGPAPFNHFYEQN-BRIANGS 253
 Db 189 LAGRDAYIFLELFFAFPHL-----RSNDFHAGESYAGHYIPQIAHETVVKNER-----239
 QY 254 VNGVQLNFSIGIINGIIDEAIOAPYPPEFAV-NTYGIKAVNETVYNYMKFANQMPNGC 312
 Db 240 ----TFNLTSVMIGNITDPLIQADYBPMACGKGYPVLSSECEKWSKAAGR-----290
 QY 313 QDLISTCKQNTALADYAL--CAEATNMCRDNVEGYPYAFAGRGVYDIRHPYDDP-----366
 Db 291 -----CARLNKLCVASKSSLPICVATAYCDASALLEPYIN-TGLNVYDIRGPCEDNSTG 343
 QY 367 ---TPPSYINKFLAKDSVMDAIGVNI-NTYQSNNDVYVAFQOTGDFVWPNFIEDLEETLA 422
 Db 344 MCYTGLRVQVMYMFPEVOETLGSVDVNYSCDNDVFTGFLFTGDSKP-FQQYIAELLN 402
 QY 423 LPVTVSLIYGADYICNWFQGOAVSLAANYSOAAQFRSAGYTP--LKVNGVVEYGETREYG 480
 Db 403 HNIPLVIYAGDKYICNWLGNHWSNELEWINKRYQRMLRPWVSKETGELGOVKNYG 462
 QY 481 NFSFTRVVEAGHEVPYQYPTASLOLFNRTIPG 512
 Db 463 PFTFLRYDAGHMVYDQPEASLEMVNSWISG 494

RESULT 4

US-10-369-493-13233
 ; Sequence 13233, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; CURRENT FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; PRIOR FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 13233
 ; LENGTH: 429
 ; TYPE: PRT
 ; ORGANISM: Aspergillus nidulans
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (1)..(429)
 ; OTHER INFORMATION: unsure at all Xaa locations
 US-10-369-493-13233

Query Match 18.8%; Score 559.5; DB 15; Length 429;
 Best Local Similarity 30.7%; Pred. No. 9.9e-44;
 Matches 149; Conservative 71; Mismatches 169; Indels 97; Gaps 19;
 QY 65 PGVKSYSYGVDTSP-ESHTFF-----WFFEARHNPETAPITLWNGG 105
 Db 4 PDVKQVTGLDNDENDKHLFYLVXVDIIVVIGRVYXRAEGFFESRNDKNDPVLWNGG 63
 QY 106 PGSDSLIGLFEELGPGCHVNSTFDYINPHSWNEVSNLLFLSQPLGVGFYSYTDVGSINP 165
 Db 64 PGCSLSITGLFMELGPSIDENIKPVNPFYAMNSASVIFLDQPVNVVYSYS-----114
 QY 166 VTGVVNSSFAGVQGRYPTIDATLIDTTNLAAEAAWEIILQGLSLGPSLDSRVQSKDFSL 225
 Db 115 -----GSTVSDTV-AAGKDVVALLTLFFKQFP-----EYAQDQFHI 149
 QY 226 WTESYGGHYGAFFNHFYQNERIANGSVNGVQLNFSINGIINGIIDEAIOAPYPPEFAV 285

Db 150 AGESYAGHYIIPVFTSEILSHOKR-----NINLKSVLINGLTLGLTQVEYYRPMAC 200
 QY 286 NNTYGIKAVNETVYNYMKFANQMPNG-----CODLISTCKQNTALADYALCAEATNMCRD 342
 Db 201 GEG-GYPAVLDE-----SSCRSDNALGRCSQMIESCYNSESA-----WVCVPASIYCNN 249
 QY 343 NVEGYPYAFAGRGVYDIRHPYDDPT-----PPSYNKFLAKDSVMDAIGVNI-NTYQSN 396
 Db 250 ALLAPYQR-TGQNYDVRGKCEDESNLCYKMGVYSEYINKPEVRAAAGAEVDGYDCNF 308
 QY 397 DVYVAFQOTGDFVWPN--NFTEDLEIILALPVVSLIY-GDADYICNWFQGOAVSLAANY 453
 Db 309 DINENFLFHGDMKPYHRLVPGILE--QIPV---LIYAGDADFICNMLGNKAWTEALEWP 363
 QY 454 QAAQFRSAGYTPPKV-----NGVEYGETREYGNFSFTRVVEAGHEVPYQYPTASLOLFN 508
 Db 364 GHKEFAAAPMEDLKIVDNEHTGKIGIKTHGNFTFMRLYGGGHVWPMQPEASLEFENR 423
 QY 509 TIFG-W 513
 Db 424 WLGGEW 429

RESULT 5

US-10-369-493-3911
 ; Sequence 3911, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; CURRENT FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; PRIOR FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 3911
 ; LENGTH: 539
 ; TYPE: PRT
 ; ORGANISM: Neurospora crassa
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (1)..(539)
 ; OTHER INFORMATION: unsure at all Xaa locations
 US-10-369-493-3911

Query Match 18.5%; Score 550; DB 15; Length 539;
 Best Local Similarity 30.7%; Pred. No. 1.1e-42;
 Matches 154; Conservative 66; Mismatches 187; Indels 94; Gaps 18;
 QY 47 NVTIRYKEPGAGVCTTPGVKSYGYVDTSP-ESHTFF-----84
 Db 99 NLRVKSVDPFSLGVDK-----VKQYSGYLDEENDKHLFYLVXVNHARMYLVSNKHGADSDC 154
 QY 85 -WFFEARHNPETAPITLWNGGPGSDSLIGLFEELGPGCHVNSTFDYINPHSWNEVSNLL 143
 Db 155 PGFFESRNDKNDPVLWNGGPGCSLTGLFLELGPSIDKRLKRVSNVEYAMNNASVI 214
 QY 144 FLSQLGVGFYSYTDVGSINPVTGVWENSSFAGVQGRYPTIDATLIDTTNLAAEAAWEI 203
 Db 215 FLDQPVNVVYSYS-----GNAVSTVAAGKDVVAL 244
 QY 204 LQGLSLGPSLDSRVQSKDFSLWTSYGGHYGPAPFNHFYQNERIANGSVNGVQLNFS 263
 Db 245 LTLFFHQFP-----EYAKQDFHAGESYAGHYIPVFASEILSHKDR-----NINLKS 291
 QY 264 LGILINGIIDEAIOAPYPPEFAVNNTYGIKAVNETVYNYMKFANQMPNGCQDLISTCKQTN 323

380	YPTLQD-IDDYLNQDYYKBAVGAEDVHYBESCMPFDINRNPLFAGDWMKP-YHTAVTDLLNQ 437	
424	PVEVSLIYGDADYICNMFGGQAVSLAANYSQAAQFRS---AGYTPKVNVEYGETREY 480	
438	DLPIVLVAGDKDFICNWLGNKAWTDVLPKDYDEEFASQKVRNWTASITDEVA-GEVKS 496	
481	NFSFTRVYRAGHEVPPYQPIASLQFNRTIFG 512	
497	HFTYLVFVGGHVPFDPENALSMVNEWIHG 528	
<p>RESULT 7</p> <p>US-09-420-785A-4</p> <p>; Sequence 4, Application US/09420785A</p> <p>; Patent No. US20010010923A1</p> <p>; GENERAL INFORMATION:</p> <p>; APPLICANT: MORTENSEN, UFFE</p> <p>; APPLICANT: OLESEN, KOELD</p> <p>; APPLICANT: STENNICKE, HENNING</p> <p>; APPLICANT: SORENGSEN, STEEN B.</p> <p>; APPLICANT: BREDDAM, KLAUS</p> <p>; TITLE OF INVENTION: MODIFIED CARBOXYPEPTIDASE</p> <p>; FILE REFERENCE: 089187/0109</p> <p>; CURRENT APPLICATION NUMBER: US/09/420,785A</p> <p>; CURRENT FILING DATE: 1999-10-19</p> <p>; NUMBER OF SEQ ID NOS: 4</p> <p>; SOFTWARE: PatentIn Ver. 2.1</p> <p>; SEQ ID NO 4</p> <p>; LENGTH: 421</p> <p>; TYPE: PRT</p> <p>; ORGANISM: <i>Saccharomyces cerevisiae</i></p> <p>US-09-420-785A-4</p>		
<p>Query Match 17.4%; Score 517.5; DB 9; Length 421;</p> <p>Best Local Similarity 29.5%; Pred. No. 8.9e-40;</p> <p>Matches 142; Conservative 70; Mismatches 185; Indels 85; Gaps 17;</p>		
51	RYKEFGAEGVCETTPGVKSYGYVDTSPE-SHTFFWFPEARHNPETAPITLWNGPGSD 109	
1	KIKDPKILGI--DPNVTVQTYGLDVEDEDKHFFFTFESRNDPAKDPVILWNGPGCS 57	
110	SLIGLFEELGCHVNSTFDDYINPHSNWVSNLLFLSQPLGVGFSYSDTVGDSINPVTG 169	
58	SUTGLFELGPGSSIGDPDLKPIGNPYSNWNSNATVIFLDQPVNVGFSYS-----104	
170	VENSFAGVQGRYPTIDATLIDTTNLAAEAAWEILQGLSLPSLDSRVQSKDFSLWTES 229	
105	-----GSSG-----VSNTVAAGKDVYNLELFFDQFPEYVKNQO--DFHIAGES 146	
230	YGGHYGPAFFNHFYEQNERIANGSVNGVQLNFNGLII--NGIIDEALQAPYYPEFA---284	
147	YAGHYIPVFASEILSHKDR-----NFNLTSVLIGNGLTDPLOTQNYVEPMAAGE 195	
285	-----VNNTYGIKAVNEVYVNMKPFANQMPNGCODLISTCKQNRRTALADYALCAEATNM 339	
196	GGEPSVLPSEECASMEDSL-----ERCLGLTESCYDSQ-----SWSCVPATII 239	
340	CRDNVEGPPYAFAGVGYDIRHPYDD-----PTPPSYNNKFLAKDSVMDALGVNINTQS 394	
240	CNNAQLAPYQR-TGRNVYDIRKDCGGNLCYPTLQD-IDDYLYNQDYVKEAVGAEDVHYES 297	
395	-NNDVYVAFQQTGDFWPNFIEDLEIILALPVRSLLIYGDADYICNMFGGQAVSLAANY 453	
298	CNFDINRNPLFAGDWMKP-YHTAVTDLLNQDPIILVYAGDKDFICNWLGNKAWTDVLPWK 356	
454	QAAQFRS---AGYTPKVNVEYGETREYGNFSTRVYVYAGHEVPPYQPIASLQFNRTI 510	
357	YDEEFASQKVRNWTASITDEVA-GEVKSCHFTYILRVFNGHGWPFDPENALSMVNEWI 415	
511	FG 512	
416	HG 417	

RESULT 8
 US-09-901-252-15
 ; Sequence 15, Application US/09901252
 ; Patent No. US20020026658A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Chapple, Clint
 ; TITLE OF INVENTION: Genes Encoding Sinapoylglucose:Malate Sinapoyltransferase and Met
 ; TITLE OF INVENTION: Use
 ; FILE REFERENCE: NL422-004
 ; CURRENT APPLICATION NUMBER: US/09/901,252
 ; CURRENT FILING DATE: 2001-07-09
 ; PRIOR APPLICATION NUMBER: US 60/216593
 ; PRIOR FILING DATE: 2000-07-07
 ; NUMBER OF SEQ ID NOS: 17
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 15
 ; LENGTH: 421
 ; TYPE: PRT
 ; ORGANISM: Saccharomyces
 US-09-901-252-15

Query Match		17.4%	Score 517.5;	DB 9;	Length 421;
Best local similarity		29.5%;	Pred. No. 8.9e+40;		
Matches 142;	Conservative	70;	Mismatches 185;	Indels	Gaps 17;
QY	51	RYPEKGAEGVCETTPGVKSISGVVDTSPE-SHTFFWFPEARHNPETAPITLWLNGGPGSD	109		
Dd	:	: :	: :	: :	: :
Dd	1	KIKDKPKILGI--DPNVQTQVGLDVEDEOKHFFFWTFESRNDPAKPVLWLNGGPGCS	57		
QY	110	SLIGLEELCPCHVNSTFDDYINPHSNWEYSNLLFLSQPLVGVSYSYDVTVDGSINPVTGV	169		
Dd	:	: :	: :	: :	: :
Dd	58	SLTGLELGPSSIGEDLPKIPGPSWNSNATVIFLDQPVNVGFSYS-----	104		
QY	170	VENSSFAGVQGRYPTIDATLIDTNLAEEAAEWELQGLSGLPSLDSRVOSKDFSLWTES	229		
Dd	:	: :	: :	: :	: :
Dd	105	-----GSSG-----VSNVAAGKDVPYNLELFFDQFEYNVKQG--DFHIAGES	146		
QY	230	YGGHYPAPFNHPHYBONERIANGSVNGVQLNFSNLGII--NGIIDEAIQAPYYPEFA---	284		
Dd	:	: :	: :	: :	: :
Dd	147	YAGHYIPVFASEILSHKDR-----NFNLTSVLIGNGLTDPLTYNYEPMACGE	195		
QY	285	-----VNNTYGIKAVNETVVNMKFANQPMPNGCQDLISTCKQNRNALADYALCAEATNM	339		
Dd	:	: :	: :	: :	: :
Dd	196	GGEPSVLPSEECAMEDSL-----ERCLGIESCYDSQ-----SWMSCVPATIY	239		
QY	340	CRDNVEGPPIYAFAGRGVYDIRHPYDD-----PPPPSYNNFKLAKDSVMDAIGVINYTQS	394		
Dd	:	: :	: :	: :	: :
Dd	240	CNNAQLAPYQR-TGRNVYDIKDCCEGNLCYPTLOD-IDDYLNQDVTKGAVGAEVHDYES	297		
QY	395	-NDVITYAFQQTGDFFVMNPFIEDLEBILALPVRVSLIYGDADYICNFWFGQQAVSLAANTYS	453		
Dd	:	: :	: :	: :	: :
Dd	298	CNFDINRNFLFAGDMWK-P-YHTAVTDLLNQDPLILVYAGDKDFICNWLGNKAWTDVLPWK	356		
QY	454	QAQAQFRS---AGVYTLKVNGVEYGETREYGNFSFTFYVYEAQHEVPYQYIASIQLPNRTT	510		
Dd	:	: :	: :	: :	: :
Dd	357	YDEEFASQKVRNWTAISITDEVA-GEVSKSYKHFTYLEVFNGGHMVPDPDENALSIMVNEWI	415		
QY	511	FG	512		
Dd	:				
Dd	416	HG	417		

RESULT 9
US-10-425-114-48698
; Sequence 48698, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kowalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E

```

: APPLICANT: Cao, Yongwei
:
: TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
:
: TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
:
: FILE REFERENCE: 38-21(53313)B
:
: CURRENT APPLICATION NUMBER: US/10/425,114
:
: CURRENT FILING DATE: 2003-04-28
:
: NUMBER OF SEQ ID NOS: 73128
:
: SEQ ID NO 48698
:
: LENGTH: 500
:
: TYPE: PRT
:
: ORGANISM: Glycine max
:
: FEATURE:
:
: OTHER INFORMATION: Clone ID: 700563631_FLI.pcp
US-10-425-114-48698

```

[illegible]

RESULT 10

```

US/10-425-114-44602
; Sequence 44602, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecu
; FILE OF INVENTION: Plants and Uses Th
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128

```

```
; SEQ ID NO 44602
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 701038279_FLI.pep
US-10-425-114-44602

Query Match      15.9%; Score 473; DB 12; Length 525;
Best Local Similarity 28.5%; Pred. No. 2e-35;
Matches 133; Conservative 77; Mismatches 182; Indels 74; Gaps 17;

Qy 62 ETTGCVKS---YSGY--VDTSPESHTEFWFEARHNPTAPITLWLNKGGSDSLIGLFE 116
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 103 DSEPSVEDLGHAGYPIQHSAAARMFYFFESNRKED-PVVIWLTGGPCSSSELALFY 161
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 117 ELGPGHVNSTDDYINPHSNWNEVSNLLFLSOPLGVGFSYSDTVGSGINPVTGVVENSFA 176
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 162 ENGPFKIADNLSLVNWEYGMKASNLVYDQPTGTGFSYSDLRDIRHNEGV----- 214
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 177 GVQGRYPTIDATLIDTTNLAABAAWEILQGLSGLPSLDSRVQSKDFSLWTSYGGHYGP 236
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 215 -----SNDLYDFIQAPFVEHP-----QYAKNDFFITGESYAGHYIP 250
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 237 AFFNHFYQNERIANGSVNGVLNENSLGIINGIIDEALQAPYYPEFAVNNTYGIKAYNE 296
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 251 AFATRIHRGNK-----AKEGITHINLKGALGNGLNPALQYKAYPDYALENGI-IKKATR 304
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 297 TVYNMKEFANOMPNGCDQLISTCKQTNRTA-LADYALCAEATNMCRDNVGGPYAFAG-R 354
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 305 NLLNLV-----IVPACESAIKLCGTNGKTSMAAYVVC-----NLVFSDI---MLHAGDT 351
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 355 GYDIRHP-----YDDPTPPSYNKKFLAKDSVMDAIGV-NINYTQSNNDVYAFQQTGD 407
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 352 NYDIRKKCEGSLCYDF-----SNMDKFLNQSVRDSLGVGKLFHVSCTEVYAAMLVD-- 405
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 408 FWPWFPIE-DLEEILALPVRSLIYGDADYICNWFGGQAVSLAANYSAQAQFRSAGYTPL 466
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 406 --WMRNLEVGIPDLLEDGINLLIYAGEYDLICNWLGNRWVHAMEWSGQKEFATISLEVPF 463
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 467 KYNVEYGETREYGNPSFTRVYEAGHEVYQPIASLQLFNRTIFG 512
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 464 VVDGSEAGLLKSYGLPSFLKVNAGHVMFMDQPKAALEMLKKWING 509
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 11
US-10-437-963-155055
; Sequence 155055, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazov, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 155055
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(525)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; YSGY--VDTSPESHTEFWFEARHNPTAPITLWLNKGGSDSLIGLFEELGPGCHVNSTF 127
```

```
; OTHER INFORMATION: Clone ID: PAT_MRT4530_54856C.1.pep
US-10-437-963-155055

Query Match      15.7%; Score 468; DB 16; Length 525;
Best Local Similarity 27.4%; Pred. No. 6e-35;
Matches 138; Conservative 85; Mismatches 202; Indels 78; Gaps 18;

Qy 23 TPASVGRRLQPKNPTGVKTLTANNVTIRYKEPAEGVCETTPGKYSYGY--VDTSPES 80
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 72 SPSSRGDGDVPAGTLVERPIHLASMATKSGSSAE-----DLGHHAGYVRLPNTHDA 124
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 81 HTFPWFPEARHP-ETAPITLWLNKGGSDSLIGLFEELGPGCHVNSTFDDYINPHSNVE 139
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 125 RLFPFFESRSGKGEDDPVVIWLTGTGGPCSSSELALFYENGFFHTADNMSLVWMDFGWDQE 184
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 140 SNLLFLSQPLGVGSYSDTVGSGINPVTGVVENSFAVGQGRYPTIDATLIDTTNLAEEA 199
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 185 SNLIYDQPTGTGFSYSNPRDTRHDEAG-VSNDLYA----- 220
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 200 AWEILQGLSLGSLDSRVQSKDFSLWTSYGGHYGPAFFNFHYEQNERIANGSVNGVOL 259
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 221 ---FLQAFFTEHPNF---AKNDFVITGESYAGHYIPAFASRVYKGNK-----NSEGIHI 268
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 260 NFNSLGI--NGIIDEALQAPYYPEFAVNNTYGIKAVNETVYNMKEFANOMPNGCODLIST 318
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 269 NLKGFXAIGNGLTDPALQYKAYTDYSLD---MGLITKSOQFNRI---NKIVPTCELAIKL 321
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 319 CKQINR-TALADYALCAEATNMCRDNVGGPYAFAGRVGYDIRHP-----YDDPTPPSY 371
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 322 CGTGTGTCGLGAYVVC-----NLIFSSIB---TIIGKNYVDIRKECVGSLCYD---LSN 370
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 372 YNKFLAKDSVMDAIGV-NINYTQSNNDVYAFQQTGDFVWPNFTB-DLEEILALPVRVSL 429
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 371 MEKFLQKSVRESLGVGDIQFVSCSPTYQAML---LDWMENLEVGIPELLENDIKVLI 426
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 430 IYGDADYTCNWFGGQAVSLAANYSAQAQFRSAGYTPLKVNVEYGETREYGNFSFTRYE 489
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 427 YAGEYDLICNWLGNRWVNSMEWSGKEAFVSSSEPTFVDGKGAGILKSYGLPSFLKVDH 486
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 490 AGHEVYQPIASLQLFNRTIFG 512
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 487 AGHVMFMDQPKVALEMLMRWTSG 509
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 12
US-10-424-599-202253
; Sequence 202253, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 202253
; LENGTH: 462
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_2465C.1.pep
US-10-424-599-202253

Query Match      15.5%; Score 463; DB 12; Length 462;
Best Local Similarity 27.3%; Pred. No. 1.5e-34;
Matches 124; Conservative 80; Mismatches 180; Indels 70; Gaps 15;

Qy 70 YSGY--VDTSPESHTEFWFEARHNPTAPITLWLNKGGSDSLIGLFEELGPGCHVNSTF 127
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
; YSGY--VDTSPESHTEFWFEARHNPTAPITLWLNKGGSDSLIGLFEELGPGCHVNSTF 127
```



```

54 HAGYYRLPRSKAARMFYFFFEBS-SKKNDPPVIMLTGGPGCSSELALFYENGPPQLTKNL 112
QY 128 DDYINPHSWNEVSNLELSQLGVGFSYSDTVDGSIINPVTGVWENSSFAGVQGRYPTIDA 187
Db 113 SLVWNIDYGWDKASNIIFVDQPTGTFGSYTSDESDIRHDEBVG- 154
QY 188 TLIDTNTLAAEAWEILQGLFSLGSLDSRSVQSKDFSLWTSYGGHYGPAFFNHHYEQNE 247
Db 155 -----SNDLYDFLQAFKEHP-----QFTKNDFYITGESYAGHYIIPALASRVHQGNK 201
QY 248 RIANGSVNGVOLNFNSLGIINGIIIDAIQAPYYPEFAVNNTYGIKAVNETVYNYMKFANO 307
Db 202 -----AKESIHNLKGFALGNGLTNPEIQAYTDTYALDR--GL--IRKAEVNSI---NK 249
QY 308 MPNGQODLSTC-KQTNRTALADYALCAEATNMCRDNVEGPPYAFAGRGVYDIRHP--- 362
Db 250 LIPPCQKQAEACGTEGGTCVSSLVCKNIFNRIMTIADDVNY-----YDIRKKCVGV 302
QY 363 --YDDPTPPSYNNFLAKDSVMDAIGV-NINVTQSNNDVYAFQQTGDFVMPNFIE-DLE 418
Db 303 LCYDF-----SYMEDFLNEKTVRDALGVGDLDVSCSSTVYSAMQD-----WNRNLEVGI 354
QY 419 EIALPVRVSLIYGADYICNWFGQAVSLAANTYSQAQFRSAGYTPUKVNGVEYGETRE 478
Db 355 TLLBEGIKVLVYAGEEDLICNLWGNRSRVQAMENSGQKQFGASGTVPLVLDGAEGTLKS 414
QY 479 YGNFSFTRVYEAGHEVPYYPQIASIQLPNRITFG 512
Db 415 HGPIAFULKVEAGHVPMDQPKALEMLRSWMQG 448

```

```

RESULT 13
US-10-382-239A-2
; Sequence 2, Application US/10382239A
; Publication No. US2003021951A1
; GENERAL INFORMATION:
; APPLICANT: Societe des Produits Nestle
; TITLE OF INVENTION: No. US2003021951A1el Carboxypeptidase of Cocoa
; FILE REFERENCE: 884e5-7006
; CURRENT APPLICATION NUMBER: US/10/382.239A
; CURRENT FILING DATE: 2003-03-05
; PRIOR APPLICATION NUMBER: EP 01116407.6
; PRIOR FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: PCT/EP02/07162
; PRIOR FILING DATE: 2002-06-28
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin ver. 3.1
; SEQ ID NO 2
; LENGTH: 508
; TYPE: PRT
; ORGANISM: Theobroma cacao
US-10-382-239A-2

```

```

Query Match      15.5%; Score 462.5; DB 15; Length 508;
Best Local Similarity 27.5%; Pred. No. 1.9e-34;
Matches 150; Conservative 88; Mismatches 203; Indels 105; Gaps 22;

QY      4 YEF-LGVLPVLAASWA-----LPGSTPASVGRQL-----PKNFTGVKLTLLTA 45
Db      8 YPESVLLFLISISAAASFLDERLGGSPFSIIHAKLIRELNLNLFPEEKNV---VDG 64

QY      46 NNVTI-----RYKEPGAEGVCETTGPVK-----YSGY--VDTPESHTFFWFFE 88
Db      65 GQVSLPEDSLVRKRFKNL----AVPGGVAVEDLGHGAGYKLANSHDARWFYFFE 119

QY      89 ARHNPTAPITLWLGPGSDSLIGLFEELGPGCHVNSTDDYINPNSHWEVSNLLFLSQP 148
Db      120 SR-NSKKDPVWILTGPGCSSELAFYENGPTIAENSLIWNQ'GWMASNLXVDQ' 178

QY      149 LGVGFYSYSDTVGDSINPTGVVNSFAGVQGRYPTIDATLIDTTNLAAABAAEIIQGLF 208
Db      179 IGTGFSYS-----SRRDIRHNEDEVSNLDY-----FLQAF 211

```

QY	209	SGLPSLDKRVOKDFSLWTESGHYGPAFFNFHFEQNERIANGSVNGVOLFNSLGIIN	268
Db	212	AEHPEF---EKNDFYITCESYAGHYIPAPARVHQGNK-----AKDGHINLKGPAIN	262
QY	269	GIIIDRAIQAPPYPEFAVNNTYGIKAVNETVTNYMKFANQMPCGDILSTCKQTNR-TAL	327
Db	263	GLTDPAIQVKAYTDVALD---MGVIKKSDYNRI--NKLVPVCEMAIKLCGTGGTISCM	315
QY	328	ADYALCAEATNMCRDNVEGPIYAFAG-RGVDIRHPYDDPTPPSYN--KFLIAKDSVMDA	384
Db	316	ASYFC-----NAIFTGIMALAGDTNYDIIRTKCEGSLCYDFSNNMETFLNQESVRDA	367
QY	385	IGV-NINYTSQNNVVYAFQQTGDGFVWFNFIE-DLEEIILAUPRVSLSLYGDA DYTCNWFG	442
Db	368	LVGSIIDFVSCPTVYQAWLV D----WMRNLEVGIPALLEDGVKLLVYAGEYDLICNWL G	423
QY	443	GQAVSLAANYQAQFRSAGTYPLKVNGVVEYGETREYGNFSFTRYVEAGHEVPYQP TAS	502
Db	424	NSRWVHAMEWSEQEFVASPEVPFVVDGSEAGVLRTHTGLFLKVHDAGHVMPMDQPKAA	483
QY	503	LQLFNRR	508
Db	484	LEMILKR	489

RESULT 14

```

US-10-424-599-214407
; Sequence 214407, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 214407
; LENGTH: 484
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(484)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_35637C.1.bep
; US-10-424-599-214407

```

Query Match	15.4%	Score 458.5	DB 12	Length 484
Best Local Similarity	27.5%	Pred. No. 4.2e-34		
Matches 150; Conservative	76;	Mismatches 211;	Indels 109;	Gaps 22;
QY	7	LSVLPPLVAASW-----ALPGSTPASVGRRL-----PKNPTGV-----KTLTTANN	47	
Db	5	VSLLLVSLFSYATSGFNHEHAYPQSQAEMLTRNLNLPKQPVNIIKGFDSFVPGKI	64	
QY	48	VTIRYKEPGAEGVCETTPGVKSYGVYDTSPESH-----FWFFEARHNPEATITLWL	102	
Db	65	VEKKFSLLGHSG-----PSIQHLGHAGQYSLPHSKAGSRFYFFESRNKKD--PVIWL	118	
QY	103	NGCGFSDSLGLFEEELGPCHVNSTEDDYINPHSWNEVSNLLFLSQGLGVGFSTDVWGS	162	
Db	119	TGCGFGGSELALFYENGPHIAANLSLTWNQYQMDQASNILFYDQPTGTGFSYSSD--DSD	177	
QY	163	INPVTGVWNSFPAGVQGYPTIDATLIDTNLAAEAAWEILGELSLGPLSDSRVQSKD	222	
Db	178	IRHDEASIND-----LYDFQEFKKAHP-----KFKVND	207	
QY	223	FSLWTSYSGHYGPAFNHFHYEQNERIANGSVNGVOLFNSLGIINGIIDEAIOAPYPE	282	

Db 208 FYITGESYAGHYTPAHASRIIQGNK-----ENOGIYINLKLGAIGNGATPAIQYQYPD 262
QY 283 FA-VNNTYGIKAVNETVYNYMKFANOMPNGCODLISTCKQTNRTALADYALCAEATNMC 341
Db 263 FAXLDNKKIITYA-----NDEINKLIPD-CEQANCTCETGGOS-----CAIAFNTCQ 309
QY 342 -----DNVEGPPYAFAGRGVYDIRHP-----YDDPTPEPSYNNKFLAKDSVMDAIGV- 387
Db 310 KIFVHILDFAFGINY-----YDIRKKCKGDMWCYDFRNVETLLN--LPK--VKSVIGVS 358
QY 388 -NINTQSNNDVYAFQGTGFVWPNFIEDLEEILALPVRVSLYIGDADYICNWFQGOAV 446
Db 359 NDLOVYVSCSKVHEAMQ--DYM-RNMEVEIPSLLEDGIKLVYVGBEDLLICNWLGNRW 415
QY 447 SLAANYSQAAQFRSAGYTPLKVNQVEYGETREYGNFSTRVYEAGHEVPYYPQIASLQLF 506
Db 416 VHMKWSGKKAFGKSPVKFVVDGSKAGSLNSYGLSFLKLVHEAGHLVPMQPKAALQML 475
QY 507 NRTIFG 512
Db 476 QSMWAG 481

Search completed: July 22, 2004, 21:29:56
Job time : 51 secs

345 LEKYLQKSVRDALGVGDIDFVSCSSTVYQA-----MLVDWMRNLE--VGIPALLEDG 395
QY 429 ---LIY-GDADYICNWFQGOAVSLAANYSQAAQFRSAGYTPLKVNQVEYGETREYGNFSF 484
Db 396 INMLVYAGEFDLLICNWLGNKSWHAMWESQGEFVVSSEVPFTVDDSEAGLLKKYGPLSF 455
QY 485 TRVVEAGHEVPYYPQIASLQLFNR 508
Db 456 LKVDHAGHMVPMQPKASLEMLXR 479

RESULT 15
US-10-424-599-172991
; Sequence 172991, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 172991
; LENGTH: 499
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_127226C.1.pep
US-10-424-599-172991

Query Match 15.3%; Score 455; DB 12; Length 499;
Best Local Similarity 29.1%; Pred.No.9.4e-34;
Matches 129; Conservative 68; Mismatches 175; Indels 72; Gaps 14;
QY 78 PESHT---PFWPFARHNPETAPITLWLNKGFGSDSLIGLFEELGFCVHNSTFDDYINPH 134
Db 95 PHSHAAKMEYFFESR-NSKOPVVIWLTGGPGCSSELAVFYENGPPFKIANNNSLWNEY 153
QY 135 SNNEVSNLLFLSOPLGVSYSSTVDGSIINPTGVVWNSFAGVQGRYPTIDATLLDTTN 194
Db 154 GMDKVNLLYVDQPTGTFGSYSTDKRDIRHDEGV-----188
QY 195 LAEEAAWETLQGLSLGSLDRVQSKDSLWTSYGGHYGPAPFNNHFYEQNERIANGSV 254
Db 189 --SNDLYDFLQAFFAHP-----EYVKNDFITGESYAGHYIPAFARVHGNK-----AK 237
QY 255 NGVQLNFSLGIINGHIDEAIOAPYPYEPFANNYTGKAVNETVYNYMKFANOMPNGCQD 314
Db 238 EGIHINLKGPAIGNGLTDPGQYKAYTDYALDMGIIOKA-----DYERINKVWVPACEM 291
QY 315 LISTCKQTNRTA-LADYALCAEATNMCNDRNVEGPPYAFAGRGVYDIRHPYDDPTPPSYN 373
Db 292 AIKLCGTGDKIACCTASYFVCNTIFNSIMSHAGDINY-----YDIRKCEGSLCYDFS 344
QY 374 --KFLAKDSVMDAIGV-NINTQSNNDVYAFQGTGFVWPNFIEDLEEILALPVRVS-- 428

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 29, 2004, 11:29:19 ; Search time 112 Seconds
(without alignments)
2749.980 Million cell updates/sec

Title: US-09-712-338-2
Perfect score: 2979
Sequence: 1 MRGVEFLSVLPVAAWALP.....HTQSSVPLPTATSSVGVMA 555

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09712338/runat.22072004.170321.10374/app.query.fasta.1.711
-DB=Issued Patents NA -OPM=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=spt -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09712338 @CGN 1 1 56 @runat.22072004.170321.10374 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/2/ina/5A COMB seq.*
2: /cgn2_6/ptodata/2/ina/5B COMB seq.*
3: /cgn2_6/ptodata/2/ina/6A COMB seq.*
4: /cgn2_6/ptodata/2/ina/6B COMB seq.*
5: /cgn2_6/ptodata/2/ina/6C COMB seq.*
6: /cgn2_6/ptodata/2/ina/6D COMB seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2975	99.9	1662	3	US-08-943-714-1 Sequence 1, Appli
2	623	20.9	2002	1	US-08-309-341-3 Sequence 3, Appli
3	623	20.9	2002	1	US-08-608-267-3 Sequence 3, Appli
4	623	20.9	2002	1	US-08-608-452-3 Sequence 3, Appli
5	623	20.9	2002	1	US-08-608-224-3 Sequence 3, Appli
6	623	20.9	2002	2	US-08-967-149-3 Sequence 3, Appli
7	594	19.9	2068	1	US-08-309-341-1 Sequence 1, Appli
8	594	19.9	2068	1	US-08-608-267-1 Sequence 1, Appli
9	594	19.9	2068	1	US-08-608-452-1 Sequence 1, Appli
10	594	19.9	2068	1	US-08-608-224-1 Sequence 1, Appli
11	594	19.9	2068	2	US-08-967-149-1 Sequence 1, Appli
12	568.5	19.1	2503	1	US-09-640-305-3 Sequence 3, Appli

13	568.5	19.1	2503	1	US-08-360-673-3 Sequence 3, Appli
14	523.5	17.6	2632	2	US-08-899-324-32 Sequence 32, Appli
15	523.5	17.6	2632	3	US-08-329-892B-32 Sequence 32, Appli
16	361	12.1	1551	2	US-08-828-488-4 Sequence 4, Appli
17	361	12.1	1551	4	US-09-299-689A-4 Sequence 4, Appli
18	359	12.1	2076	4	US-09-907-794A-163 Sequence 163, App
19	359	12.1	2076	4	US-09-907-794A-163 Sequence 163, App
20	359	12.1	2076	4	US-09-907-794A-163 Sequence 163, App
21	348.5	11.7	1670	2	US-08-828-488-2 Sequence 2, Appli
22	348.5	11.7	1670	4	US-09-299-689A-2 Sequence 2, Appli
23	339.5	11.4	1814	4	US-09-702-705-319 Sequence 319, App
24	339.5	11.4	1814	4	US-09-736-457-319 Sequence 319, App
25	339.5	11.4	1814	4	US-09-614-124B-319 Sequence 319, App
26	339.5	11.4	1814	4	US-09-671-325-319 Sequence 319, App
27	339.5	11.4	1814	4	US-09-589-184-319 Sequence 319, App
28	275	9.2	1256	4	US-09-896-866B-6 Sequence 6, Appli
29	275	9.2	1302	4	US-09-896-866B-9 Sequence 5, Appli
30	275	9.2	1317	4	US-09-896-866B-9 Sequence 9, Appli
31	266.5	8.9	1071	4	US-09-023-655-674 Sequence 674, App
32	246.5	8.3	1604	1	US-08-665-966-9 Sequence 9, Appli
33	246.5	8.3	1604	3	US-09-041-780-9 Sequence 9, Appli
34	239.5	8.0	1650	4	US-09-907-794A-254 Sequence 254, App
35	239.5	8.0	1650	4	US-09-905-125A-254 Sequence 254, App
36	239.5	8.0	1650	4	US-09-902-775A-254 Sequence 254, App
37	169.5	5.7	1263	2	US-08-828-488-6 Sequence 6, Appli
38	169.5	5.7	1263	4	US-09-299-689A-6 Sequence 6, Appli
39	163	5.5	1960	4	US-09-280-116-133 Sequence 133, App
40	161.5	5.4	497	4	US-09-643-597-190 Sequence 190, App
41	161.5	5.4	497	4	US-09-480-884A-190 Sequence 190, App
42	161.5	5.4	497	4	US-09-542-615A-190 Sequence 190, App
43	161.5	5.4	497	4	US-09-606-421B-190 Sequence 190, App
44	154	5.2	357	4	US-09-280-116-131 Sequence 131, App
45	154	5.2	607	4	US-09-280-116-130 Sequence 130, App

ALIGNMENTS

RESULT 1
US-08-943-714-1
; Sequence 1, Application US/08943714
; Patent No. 6187578
; GENERAL INFORMATION:
; APPLICANT: Blinkovsky, Alexander
; APPLICANT: Berk, Randy
; APPLICANT: Rey, Michael
; APPLICANT: Golightly, Elizabeth
; APPLICANT: Klotz, Alan
; APPLICANT: Mathisen, Thomas Erik
; APPLICANT: Dammann, Claus
; TITLE OF INVENTION: Carboxypeptidases And Nucleic Acids
; TITLE OF INVENTION: Encoding Same
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6187578o No. 6187578disk of No. 6187578th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174

COMPUTER READABLE FORM:

COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/943,714
FILING DATE: 03-OCT-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 4990.200-US
TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1662 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 1...1662
OTHER INFORMATION:
US-08-943-714-1

Alignment Scores:
Pred. No.: 1, 21e-312 Length: 1662
Score: 2975.00 Matches: 554
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.87% Indels: 0
DB: 3 Gaps: 0

US-09-712-338-2 (1-555) x US-08-943-714-1 (1-1662)

QY 1 MetArgGlyTyrGluPheLeuSerValLeuProLeuValAlaAlaSerTrpAlaLeuPro 20
Db 1 ATGGTGGCTACGAAATTTCTCAGTGTACCTTTGGTGGAGCCAGTTGGCCCTTCCA 60
QY 21 GlySerThrProAlaSerValGlyArgArgGlnLeuProLysAsnProThrGlyVallys 40
Db 61 GGAAGTACACCGCGCTCGGTAGAACAGACAGCTACCCAAAGAACCCACCGGGGTCAAG 120
QY 41 ThrLeuThrAlaAsnAsnValThrIleArgTyrLysGluProGlyValadgluVal 60
Db 121 ACTCTTACAAACCGCAACAATGTCCATCCCGTACAGGAACCCCGGGGAGGGCGTC 180
QY 61 CysGluThrThrProGlyVallysSerTyrSerGlyTyrValAspThrSerProGluSer 80
Db 181 TCGAGAGTACCCCGGTCTCAAAATCTACTCTGGATATGTGCACACCTCTCCCGAGTCC 240
QY 81 HisThrPheThrPheGluAlaArgHisAsnProGluThrAlaProIleThrLeu 100
Db 241 CATACCTCTCTGGTCTTCGAAGCAGACACATAACCCAGAACTGCACCTATACATGG 300
QY 101 TrpLeuAsnGlyGlyProGlySerAspSerLeuLeGlyLeuPheGluGluLeuGlyPro 120
Db 301 TGGTTGAATGGTGGCCCTGGAGCGATTCTTTGATCGGTCTCTTGAAGAGTTGGGCCCT 360
QY 121 CysHisValAsnSerThrPheAspAspTyrIleAsnProHisSerTrpAsnGluValSer 140
Db 361 TGCCATGTCAATTCGACTTTTGATGACTATACATCAACCTCTCTCGTGGAAACGAGGTCTCC 420
QY 141 AsnLeuPheLeuSerGlnProLeuGlyValGlyPheSerTyrSerAspThrValAsp 160
Db 421 AATTTACTATTCCTGCTCCAGCCATTCGGAGTCGGCTTTTCATATAGTATACGGTGTAT 480
QY 161 GlySerIleAsnProValThrGlyValValGluAsnSerSerPheAlaGlyValGlnGly 180
Db 481 GGGTCCATTAAACCTGTAACTGGGGTCTGTCGAAATTCGAGCTTTTCGAGGAGTTCAAGGC 540
QY 181 ArgTyrProThrIleAspAlaThrLeuIleAspThrThrAsnLeuAlaGluAlaAla 200
Db 541 CGGTACCCCAACCATTCATGCACTCTGATCGATACCTACCAATCTTCGCGCAGAGCGGCT 600
QY 201 TrpGluIleLeuGlnGlyPheLeuSerGlyLeuProSerLeuAspSerArgValGlnSer 220
Db 601 TGGGAGATCTCTGAAGGATTCCTTAGTGAGTACCTAGCTTGGACTCTAGGGTGCAGTCT 660
QY 221 LysAspPheSerLeuTrpThrGluSerTyrGlyHisTyrGlyProAlaPheAsn 240
Db 661 AAGGACTTCACTATCGACGGAGAGCTATGGAGGGCACTATGGTCTCTGCATCTTCAAT 720

241 HisPheTyrGluGlnAsnGluArgIleAlaAsnGlySerValAsnGlyValGlnLeuAsn 260
721 CATTTTACGAGCGAATGAGAGAAATGCCAACGGTAGTGTAAATGGTGTTCAGCTTAAT 780
261 PheAsnSerLeuGlyIleIleAsnGlyIleIleAspGluAlaIleGlnAlaProTyrTyr 280
781 TTCAACTCTCTGGGAATTAATAACGGCATCATCGACGAGCGATCCAGGCCCCCTTACTAC 840
281 ProGluPheAlaValAsnAsnThrTyrGlyIleIleValValAsnGluThrValTyrAsn 300
841 CCTGAATTCGCTGTGAACAATACCTACGGTATCAAGGCTGTCAAGGCTGTCAAGGAGCCGTCTACAC 900
301 TyrMetLysPheAlaAsnGlnMetProAsnGlyCysGlnAspLeuIleSerThrCysLys 320
901 TACATGAAGTTTGCACAAACCAATGCTGATGGTTCACAGGATTTGATTTCCACCTGCAAA 960
321 GlnThrAsnArgThrAlaLeuAlaAspTyrAlaLeuCysAlaGluAlaThrAsnMetCys 340
961 CAGACAAACCGCACCGCATTAGCTGACTAGCCCTCTGCGCCGAGGACCAACATGTGC 1020
341 ArgAspAsnValGluGlyProTyrTyrAlaPheAlaGlyArgGlyValTyrAspIleArg 360
1021 AGGACCAATGTGAGGGGCCATACCTACCCCTTTGCTGGTGTGTGTGTATGATATTCGG 1080
361 HisProTyrAspAspProThrProProSerTyrTyrAsnLysPheLeuAlaLysAspSer 380
1081 CATCATATGATGACCCGACTCCCGCAAGTTATTACAACAATTTCTGGCAAGACACTCT 1140
381 ValMetAspAlaIleGlyValAsnIleAsnTyrThrGlnSerAsnAspValTyrTyr 400
1141 GTCATGAGCGCTATCGCGCTCAACATCAACTACACCCAGTCCCAATAATAGCTCTACTAC 1200
401 AlaPheGlnGlnThrGlyAspPheValTrpProAsnPheIleGluAspLeuGluIle 420
1201 GCTTTCCAGCAACACGAGCGACTTTGCTGGGCCCACTTCATCGAAGACCTCGAGGAGATC 1260
421 LeuAlaLeuProValArgValSerLeuIleTyrGlyAspAlaAspTyrIleCysAsnTrp 440
1261 CTGTGCTCTCCCGTGGCTCTCCCTCATCTATGCGCAGCGGATTCATCTGCAACTGG 1320
441 PheGlyGlyGlnAlaValSerLeuAlaAsnTyrSerGlnAlaAlaGlnPheArgSer 460
1321 TTCGCGGTTCAGCGGTTCCTCTCGCTCGAATACTCCCAAGCCGCGCAGTTCGGAAGC 1380
461 AlaGlyTyrThrProLeuLysValAsnGlyValGluTyrGlyGluThrArgGluTyrGly 480
1381 GCAGGGTACACCCCTGAAAGTCAACGCGCTCGAGTATGGGCAAACTCGCAGTATGCT 1440
481 AsnPheSerPheThrArgValTyrGluAlaGlyHisGluValProTyrTyrGlnProIle 500
1441 AATTTCTCTTCCTGCTGCTATGAGCGAGCCCATGAAGTCCCATCTACAGGCCCATC 1500
501 AlaSerLeuGlnLeuPheAsnArgThrIlePheGlyTyrAspIleAlaGluGlyGlnLys 520
1501 GCCTCCTCCTGCAATTTGTTAACCGGACTATCTTCGTTGGGATATATCGAGGGCCAGAAG 1560
521 LysIleTrpProSerTyrLysThrAsnGlyThrAlaThrAlaThrHisThrGlnSerSer 540
1561 AAGATCTGCCCGCAGCTTCAAGAGATGGAACGGCTACAGCTACGATACCATACAGTCTGTC 1620
541 ValProLeuProThrAlaThrSerMetSerSerValGlyMet 554
1621 GTGCGCTGCTACGCTACCGCTACCGAGATGTCCAGTGTGGTATG 1662

RESULT 2
US-08-309-341-3
Sequence 3, Application US/08309341
Patent No. 5594119
GENERAL INFORMATION:
APPLICANT: Xaver, Debbie Sue
APPLICANT: Thompson, Sheryl Ann
TITLE OF INVENTION: GENE ENCODING CARBOXYPEPTIDASE OF ASPERGILLUS NIGER

; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: NO. 55941190 No. 5594119disk of No. 5594119th America, Inc.
 ; STREET: 405 Lexington Avenue, Suite 6400
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 10174-6401
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/309,341
 ; FILING DATE: 16-SEP-1994
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Lowney, Karen A.
 ; REGISTRATION NUMBER: 31,274
 ; REFERENCE/DOCKET NUMBER: 4247.000-US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 212 867 0123
 ; TELEFAX: 212 867 0298
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2002 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; ORIGINAL SOURCE:
 ; ORGANISM: Aspergillus niger
 ; FEATURE:
 ; NAME/KEY: intron
 ; LOCATION: 349..411
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: join (348..412)
 ; US-08-309-341-3

Alignment Scores:
 Pred. No.: 1,9e-57
 Score: 623.00
 Percent Similarity: 48.25%
 Best Local Similarity: 31.96%
 Query Match: 20.91%
 DB: 1
 Gaps: 18

US-09-712-338-2 (1-555) x US-08-309-341-3 (1-2002)

QY 45 AlaAsnValThrIleArgTyrLysGluProGlyAlaGluGlyValCysGluThrThr 64
 DB 533 GCCTATGACTCAGGCTCAGAGACCGGATCTCTCTGGCATC-----GAC 583
 QY 65 ProGlyValLysSerTyrSerGlyTyrValAspThrSerPro---GluSerHisThrPhe 83
 DB 584 CTGGCGTAAAGCAGTACACCGGTTATCGATGACACGAGACGACGACATCTGTTTC 643
 QY 84 PheTyrPheGluAlaArgHisAsnProGluThrAlaProIleThrLeuTyrLeuAsn 103
 DB 644 TACTGTTCTTCGAGTCTCGCAATGACCGGATGAGATGACCTGTTCTCTGGCTGAC 703
 QY 104 GlyGlyProGlySerAspSerLeuIleGlyLeuPheGluGluGlyProCysHisVal 123
 DB 704 GGTGGCCCTGGATGCTCTCCCTACCGGTCCTTTTCATGAGCTGGCCCTAGCAGCATC 763
 QY 124 AsnSerThrPheAspAspTyrIleAsnProHisSerTyrAsnGluValSerAsnLeuLeu 143
 DB 764 AACAAAGACATCCAGCGGCTCTACACGACTAGCTTGAATCCMACCGCTCCGTTGATC 823
 QY 144 PheLeuSerGlnProLeuGlyValGlyPheSerTyrSerAspThrValAspGlySerIle 163

Db 824 TTCCTTGACAGCCTGTCAACGTCGGTTACTCTCTTACAGCAACTCTGCT----- 871
 QY 164 AsnProValThrGlyValValGluAsnSerSerPheAlaGlyValGlnGlyArgTyrPro 183
 Db 871 ----- 871
 QY 184 ThrIleAspAlaThrLeuIleAspThrThrAsnLeuAlaGluAlaAlaTlpGluIle 203
 Db 872 -----GTCAAGCAGACACCGTGTCTGTGCAAGGACGTCTATGCCCTTG 913
 QY 204 LeuGlnGlyPheLeuSerGlyLeuProSerLeuAspSerArgValGlnSerLysAspPhe 223
 Db 914 CTTACCCCTCTCTTCAACAACATTCCCC-----GAGTATGCCAAGCAGGACTTC 961
 QY 224 SerLeuTyrThrGluSerTyrGlyHisTyrGlyProAlaPhePheAsnHisPheTyr 243
 Db 962 CACATTGCGGTGAATCTATGCTGTCATATATCCCCGCTCTTGTTCGGAGATTTTG 1021
 QY 244 GluGlnAsnGluArgIleAlaAsnGlySerValAsnGlyValGlnLeuAsnPheAsnSer 263
 Db 1022 TCTCAAGAAAGCGC-----AACATCAACCTCGAGTCC 1054
 QY 264 LeuGlyIleIleAsnGlyIleIleAspGluAlaIleGlnAlaProTyrTyrProGluPhe 283
 Db 1055 GTTCTTATGCAACGGTCTCACCGAGCGTCTCACTCAGTACGAGTACTACCGTCCCATG 1114
 QY 284 AlaValAsnAsnThrTyrGlyIleLysAlaValAsnGluThrValTyrAsnTyrMetLys 303
 Db 1115 GCCTGTGTGTCACGCT---GGTTACCCAGCTGCTTGAGCAG---GGCTCTGCCAGGCC 1168
 QY 304 PheAlaAsnGlnMetProAsnGlyCysGlnAspLeuSerThrCysLysGlnThrAsn 323
 Db 1169 ATGGACAAACGCCCTCTCTCGC---TGCCAGTCTATGATTGAGTCTTGTCTAGTTCGAG 1225
 QY 324 ArgThrAlaLeuAlaAspTyrAlaLeuCysAlaGluAlaThrAsnMetCysArgAspAsn 343
 Db 1226 AGCGCT-----TGGGTTGTGTCGCCGCTCTCTACTACTTAACAACGCC 1270
 QY 344 ValGluGlyProTyrTyrAlaPheAlaGlyValTyrAspIleArgHisProTyr 363
 Db 1271 CTCCTTGCCCTTACCAGCGC---ACCGGACAGAAAGTCTACGATGTCTGGTAAAGTGC 1327
 QY 364 AspAspPro-----ThrProProSerTyrTyrAsnLysPheLeuAlaLys 378
 Db 1328 GAGGATAGCTCCAACTCTGCTACTCGGCCATGGGTACGTACGAGTACCTCTGAACAAG 1387
 QY 379 AspSerValMetAspAlaIleGlyValAsnIleAsn---TyrThrGlnSerAsnAsnAsp 397
 Db 1388 ACGGAGTCTATGAGGCTGTGGCGTGGAGTCAACGGCTACGACTCGTGCAACTTTGAC 1447
 QY 398 ValTyrTyrAlaPheGlnGlnThrGlyAspPheValTyrPro-----AsnPheIleGlu 415
 Db 1448 ATCAACCGCAACTCTCTCTCCACGGTGAAGTGAAGCCCTACACCGTCTGCTGCG 1507
 QY 416 AspLeuGluGluIleLeuAlaLeuProValArgValSerLeuIleTyr---GlyAspAla 434
 Db 1508 GGACTCTCTGGAG-----CAGATCCCTGTC-----CTGATCTACGCTGGTGAAGCC 1552
 QY 435 AspTyrIleCysAsnTyrPheGlyGlnAlaValSerLeuAlaAlaAsnTyrSerGln 454
 Db 1553 GATTTCATCTGCACTGGCTGGGCAACAGGCTCGACTGAAGCCCTTGAGTGGCCCGGA 1612
 QY 455 AlaAlaGlnPheArgSerAlaGlyTyrThrProLeuLysVal-----Asn 469
 Db 1613 CAGGCTGAATAGCTCCGCTAAGCTGGAGGACCTGGTGGTGGAGTAAGGACAGCAG 1672
 QY 470 GlyValGluTyrGlyThrArgGluTyrGlyAsnPheSerPheThrArgValTyrGlu 489
 Db 1673 GCAAGAAAGATCGGCGCAGGTCAAGTCCCATCCCTTCACTTCCATGCTCTCTATGCG 1732
 QY 490 AlaGlyHisGluValProTyrTyrGlnProIleAlaSerLeuGlnLeuPheAsnArgThr 509
 Db 1733 GGTGGCCCATGGTCCCGATGGACCAACCCGAGTGTGAGTCTTGAATTTCTCAACCGCTGG 1792

510 IlePheGly---Trp 513
1793 TTGGAGGTGAATGG 1807

RESULT 3
US-08-608-267-3
Sequence 3, Application US/08608267
Patent No. 5688663
GENERAL INFORMATION:
APPLICANT: Yaver, Debbie Sue
APPLICANT: Thompson, Sheryl Ann
TITLE OF INVENTION: GENE ENCODING CARBOXYPEPTIDASE OF ASPERGILLUS NIGER
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: NO. 56886630 No. 5688663disk of No. 5688663th America, Inc.
STREET: 405 Lexington Avenue, Suite 6400
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10174-6401

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/608,267
FILING DATE: 28-FEB-1996
CLASSIFICATION: 435
PRIOR APPLICATION NUMBER: US 08/309,341
APPLICATION DATE: 20-SEP-1994
FILING DATE: 20-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lowney, Karen A.
REGISTRATION NUMBER: 31,274
REFERENCE/DOCKET NUMBER: 4247,000-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 867 0123
TELEFAX: 212 867 0298
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2002 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Aspergillus niger
FEATURE:
NAME/KEY: intron
LOCATION: 349..411
FEATURE:
NAME/KEY: CDS
LOCATION: Join (348..412)
US-08-608-267-3

Alignment Scores:
Pred. No.: 1,9e-57 Length: 2002
Score: 623.00 Matches: 155
Percent Similarity: 48.25% Conservative: 79
Best Local Similarity: 31.96% Mismatches: 175
Query Match: 20.91% Indels: 76
DB: 1 Gaps: 18

US-09-712-338-2 (1-555) x US-08-608-267-3 (1-2002)

QY 45 AlaAsnValThrIleArgTyrLysGluProGlyAlaGluGlyValCysGluThrThr 64
533 GCCTATGATCTCAGGGTCAAGAGACCGATCTAGTCTCTTGGCAATC-----GAC 583
QY 65 ProGlyValLysSerTyrSerGlyTyrValAspThrSerPro---GluSerHisThrPhe 83

Db 584 CTTGGCGTAAAGCAGTACACCGGTTATCTCGATGACACAGAGAACGACAGCATCTGTTTC 643
QY 84 PheTrpPhePheGluAlaAargHisAsnProGluThrAlaProIleThrLeuTrpLeuAsn 103
Db 644 TACTGGTTCTTCGAGTCTGCAATGATGACCCGAGAAATGACCCCTGTTGTTCTTGTCGTGAC 703
QY 104 GlyGlyProGlySerAspSerLeuIleGlyLeuPheGluGluLeuGlyProCysHisVal 123
Db 704 GGTGGCCCTGGATGCTCTTCCCTCACGGTCTTTTCATGGAGCTCGGCCCTAGCAGCATC 763
QY 124 AsnSerThrPheAspAspTyrIleAsnProHisSerTrpAsnGluValSerAsnLeu 143
Db 764 AACAAAGATCATCAGCGGCTCTACAACGACTCTGGAACCTCCAAACGGCTCGGTGATC 823
QY 144 PheLeuSerGlnProLeuGlyValGlyPheSerTyrSerAspThrValAspGlySerIle 163
Db 824 TTCCTTGACGAGCCTGTCAACGTCGGTTACTCTTACAGCAACTCTGCT----- 871
QY 164 AsnProValThrGlyValValGluAsnSerSerPheAlaGlyValGlnGlyArgTyrPro 183
Db 871 ----- 871
QY 184 ThrIleAspAlaThrLeuIleAspThrThrAsnLeuAlaGluAlaAalaTrpGluIle 203
Db 872 -----GTACGACGACACCGGTTGCTGTCGCAAGGAGCTCTATGCCCTG 913
QY 204 LeuGlnGlyPheLeuSerGlyLeuProSerLeuAspSerArgValGlnSerLysAspPhe 223
Db 914 CTTACCCCTCTTCTCAACAATTCCTCC-----GAGTATGCCAAGCAGCATTC 961
QY 224 SerLeuTrpThrGluSerTyrGlyHisTyrGlyProAlaPhePheAsnHisPheTyr 243
Db 962 CACATTGCGCGTGAATCTATGCTGTCCTATATATCCCGCTCTTGTCTCGAGATTG 1021
QY 244 GluGlnAsnGluArgIleAlaAsnGlySerValAsnGlyValGlnLeuAsnPheAsnSer 263
Db 1022 TCTCACAGAGAGCGC-----AACATCACTGACGATCC 1054
QY 264 LeuGlyIleIleAsnGlyIleAspGluAlaIleGlnAlaProTyrTyrProGluPhe 283
Db 1055 GTTCTTATGGCAACGGTCTCACCAGCGTCTCCTCAGTACGAGTACTACCGTCCCATG 1114
QY 284 AlaValAsnAsnThrTyrGlyIleLysAlaValAsnGluThrValTyrAsnTyrMetLys 303
Db 1115 GCCTGTGTGACGGT---GGTTACCGAGCTCTTGGACGAG---GGCTCCTGCCAGGCC 1168
QY 304 PheAlaAsnGlnMetProAsnGlyCysGlnAspLeuIleSerThrCysLysGlnThrAsn 323
Db 1169 ATGGACAACGGCTTCTCCGCG---TGCCAGTCTATGATGAGTCTTGTCTATAGTCCGAG 1225
QY 324 ArgThrAlaLeuAlaAspTyrAlaLeuCysAlaGluAlaThrAsnMetCysArgAspAsn 343
Db 1226 AGCGCT-----TGGGTTGTGTCGCGGCTCCATCTACTGTAAACAAGCC 1270
QY 344 ValGluGlyProTyrTralaPheAlaGlyArgGlyValTyrAspIleArgHisProTyr 363
Db 1271 CTCCTTGCCCTTACCAGCGC---ACCGGACAGACGCTACGATGTTCTGCTGAAGTGC 1327
QY 364 AspAspPro-----ThrProSerTyrTyrAsnLysPheLeuAlaLys 378
Db 1328 GAGGATAGCTCCAAACCTCTGCTACTCGGCCCTGCGGCTAGCTCAGCGACTACCTGACCAAG 1387
QY 379 AspSerValMetAspAlaIleGlyValAsnIleAsn---TyrThrGlnSerAsnAsnAsp 397
Db 1388 ACCGAGGTCAATTGAGGCTGTGGCGCTGAGGTCAACGGCTACGACTCGGTGCAACTTTGAC 1447
QY 398 ValTyrTralaPheGlnGlnThrGlyAspPheValTrpPro-----AsnPheIleGlu 415
Db 1448 ATCAACCGCAACTTCTCTTCCACGGTGAATGATGAAGCCCTTACCACCGCTCTCGTTCG 1507
QY 416 AsnLeuGluGluIleLeuAlaLeuProValArgValSerLeuIleTyr---GlyAspAla 434
Db 1508 GGACTCTCTGAG-----CAGATCCCTGTC-----CTGATCACTGCTGTCAGGCC 1552


```
QY 364 AspAspPro-----ThrProSerTyrTyrAsnLysPheLeuAlaLys 378
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1328 GAGGATAGTCCAACTCTGCTACTCGGCATGGCTACGTCAGCGACTACCTGAACAAG 1387
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 379 AspSerValMetAspAlaIleGlyValAsnIleAsn--TyrThrGlnSerAsnAsp 397
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1388 ACCGAGTCATTGAGGTGTGGCGCTGAGGTCAACGCTACGACTCGTGCACATTGAC 1447
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 398 ValTyrTyrAlaPheGlnThrGlyAspPheValTyrPro-----AsnPheIleGlu 415
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1448 ATCAACCGCACTTCCTCTCCAGCGTGCTGATGAGCCCTACCAACCGTCTCGTTCG 1507
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 416 AspLeuGluGluLeuAlaLeuProValArgValSerLeuIleTyr---GlyAspAla 434
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1508 GGACTCTCTGGAG-----CAGATCCCTGTC-----CTGATCTACGCTGGTGACGCC 1552
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 435 AspTyrIleCysAsnTyrPheGlyGlyGlnAlaValSerLeuAlaAlaAsnTyrSerGln 454
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1553 GATTTCATCTGCAACTGCTGGGCAACAAGCGCTGAGCTGAAGCCCTTGAGTGCCCGGA 1612
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 455 AlaAlaGlnPheArgSerAlaGlyTyrThrProLeuLysVal-----Asn 469
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1613 CAGCTGATATGCTCGCTAAGCTGGAGGACCTGGTCTGGTTCGAGATGACACAAG 1672
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 470 GlyValGluTyrGlyGluThrArgGluTyrGlyAsnPheSerPheThrArgValTyrGlu 489
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1673 GGCAAGAAGATCGGCGCAGCTCAAGTCCCATGGCAACTTCACCTTCATGCGTCTATGGC 1732
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 490 AlaGlyHisGluValProTyrTyrGlnProIleAlaSerLeuGlnLeuPheAsnArgThr 509
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1733 GGTGGCCACATGTCCTCCGATGACCAACCGAGTCGAGTCTTGATTTCTCAACCGCTGG 1792
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 510 IlePheGly---Tyr 513
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1793 TTGGGAGTGAATGG 1807
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 5
US-08-608-224-3
; Sequence 3, Application US/08608224
; Patent No. 5705376
; GENERAL INFORMATION:
; APPLICANT: Yavex, Debbie Sue
; APPLICANT: Thompson, Sheryl Ann
; TITLE OF INVENTION: GENE ENCODING CARBOXYPEPTIDASE OF ASPERGILLUS NIGER
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESS: No. 5705376 No. 5705376disk of No. 5705376th America, Inc.
; STREET: 405 Lexington Avenue, Suite 6400
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/608,224
; FILING DATE: 28-FEB-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/309,341
; FILING DATE: 20-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowney, Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 4247,000-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 867 0123
; TELEFAX: 212 867 0298
; INFORMATION FOR SEQ ID NO: 3:
```

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2002 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: Aspergillus niger
; FEATURE:
; NAME/KEY: intron
; LOCATION: 349..411
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join (348..412)
; US-08-608-224-3
Alignment Scores:
Pred. No.: 1.9e-57 Length: 2002
Score: 623.00 Matches: 155
Percent Similarity: 48.25% Conservative: 79
Best Local Similarity: 31.96% Mismatches: 175
Query Match: 20.91% Indels: 76
DB: 1 Gaps: 18
US-09-712-338-2 (1-555) x US-08-608-224-3 (1-2002)
QY 45 AlaAsnAsnValThrIleArgTyrLysGluProGlyAlaGluGlyValCysGluThrThr 64
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 533 GCCTATGATCTCAGGGTCAGAGAACCGATCCTAGCTCTCTTGGCATC-----GAC 583
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 65 ProGlyValLysSerTyrSerGlyTyrValAspThrSerPro---GluSerHisThrPhe 83
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 584 CCTGCGTAAGACGAGTACACCGTTATCTCGATCACACGAGAAACGACAGCATCTGTTTC 643
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 84 PheTrpPheGluAlaArgHisAsnProGluThrAlaProIleThrLeuTrpLeuAsn 103
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 644 TACTGGTTCTTCGAGTCTCGCAATGACCCCGAAGTAATACCTGTTGTTCTGTGGCTGAAC 703
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 104 GlyGlyProGlySerAspSerLeuIleGlyLeuPheGluGluLeuGlyProCysHisVal 123
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 704 GGTGGCCCTGGATGCTCTTCCCTCACCGTCTTTTCATGGAGCTCGGCCCTTAGCAGCATC 763
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 124 AsnSerThrPheAspAspTyrIleAsnProHisSerTrpAsnGluValSerAsnLeuLeu 143
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 764 AACAGAAGATCCAGCGGTCTACACAGCATACGCTTGGAACTCCCAACGCGTCCGTGATC 823
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 144 PheLeuSerGlnProLeuGlyValGlyPheSerTyrSerAspThrValAspGlySerIle 163
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 824 TTCCTTGACCAACGCTGTCAACGTCGGTTACTCTTACAGCAACTCTGCT-----871
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 164 AsnProValThrGlyValValGluAsnSerSerPheAlaGlyValGlnGlyArgTyrPro 183
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 871 -----871
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 184 ThrIleAspAlaThrLeuIleAspThrThrAsnLeuAlaGluAlaAlaTrpGluIle 203
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 872 -----GTCAGCGACACCCGTTGCTGTGGCAGAGCATCTATGCCCTTG 913
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 204 LeuGlnGlyPheLeuSerGlyLeuProSerLeuAspSerArgValGlnSerLysAspPhe 223
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 914 CTTACCCCTCTTCTCAACAATTCGCC-----GAGTATGCCAGCAGGACTTC 961
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 224 SerLeuTrpThrGluSerTyrGlyHisTyrGlyProAlaPheAsnHisPheTyr 243
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 962 CACATTCGCGTGATCTCTATCTGCTGCTATATATCCCGCTCTTTGCTCGGAGATTTTG 1021
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 244 GluGlnAsnGluArgIleAlaAsnGlySerValAsnGlyValGlnLeuAsnPheAsnSer 263
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1022 TCTCAAGAGAGCGC-----AACATCAACCTGCAGTCC 1054
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 264 LeuGlyIleAlaAsnGlyIleIleAspGluAlaIleGlnAlaProTyrTyrProGluPhe 283
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1055 GTTCTTAATTGGCAACGGTCTCACCGACGGTCTCTCAGTACGAGTACTACCGTCCCATG 1114
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```



```

QY 284 AlaValAsnAsnThrTyrGlyLeuAlaValAsnGluThrValTyrAsnTyrMetLys 303
D 1115 GCCTGTGTGACGGT---GGTTACCCAGCTCTCTGGACGAG---GGCTCTCCGACGCC 1168
QY 304 PheAlaAsnGlnMetProAsnGlyCysGlnAspLeuSerThrCysLysGlnThrAsn 323
D 1169 ATGACAAACGCCCTTCCTCGC---TGCCAGTCTATGATTGAGTCTTGCTATAGTTCGAG 1225
QY 324 ArgThrAlaLeuAlaAspTyrAlaLeuCysAlaGluAlaThrAsnMetCysArgAspAsn 343
D 1226 AGCGCT-----TGGGTGTGTCCGCCCTCCATCTACTGTAACAAACGCC 1270
QY 344 ValGluGlyProTyrTyrAlaPheAlaGlyArgGlyValTyrAspIleArgHisProTyr 363
D 1271 CTCCTTCGCCCTTACCAGCGC---ACCGACAGACAGCTCTACGATGTTCTGGTAAAGTC 1327
QY 364 AspAspPro-----ThrProSerTyrTyrAsnLysPheLeuAlaLys 378
D 1328 GAGGATAGCTCCACCTCTCTACTCGGCCATGGCTACGTCAGCGACTACCTGAACAAG 1387
QY 379 AspSerValMetAspAlaIleGlyValAsnIleAsn---TyrThrGlnSerAsnAsp 397
D 1388 ACCGAGTCATGAGGCTGTGGCGCTGAGTCAACGGCTACGACTCGTGCACCTTTGAC 1447
QY 398 ValTyrTyrAlaPheGlnGlnThrGlyAspPheValTyrPro-----AsnPheIleGlu 415
D 1448 ATCAACCGCAACTCTCTCTCCACGGTGACTGGATGAAGCCCTACCAACCGCTCTGTTCCG 1507
QY 416 AspleuGluGluLeuAlaLeuProValArgValSerLeuIleTyr---GlyAspAla 434
D 1508 GGACTCTCTGAG-----CAGATCCCTGTC-----CTGATCTACGCTGTGTCAGGCC 1552
QY 435 AspTyrIleCysAsnTyrPheGlyGlyGlnAlaValSerLeuAlaAlaAsnTyrSerGln 454
D 1553 GATTCATCTGCACTGGCTGGGCAACAGCGCTGGACTGAAGCCCTTGAGTGGCCCGGA 1612
QY 455 AlaAlaGlnPheArgSerAlaGlyTyrThrProLeuLysVal-----Asn 469
D 1613 CAGGCTGAATATGCTCCGCTAAGCTGGAGGACCTGGTCTGTGTCGAGATGACACAAG 1672
QY 470 GlyValGluTyrGlyGluThrArgGluTyrGlyAsnPheSerPheThrArgValTyrGlu 489
D 1673 GGCAAGAAGATCGCCAGGTCAAAGTCCCATGCAACTTCACTTATGCTCTCTATGGC 1732
QY 490 AlaGlyHisGluValProTyrTyrGlnProIleAlaSerLeuGlnLeuPheAsnArgThr 509
D 1733 GTGGGCCACATGTTCCGATGGACCAACCCGAGTCGAGTCTTGAAATTTTCAACCGCTGG 1792
QY 510 IlePheGly---Trp 513
D 1793 TTGGGAGGTGAATGG 1807

```

RESULT 6

```

US-08-967-149-3
; Sequence 3, Application US/08967149
; Patent No. 5939305
; GENERAL INFORMATION:
; APPLICANT: Yaver, Debbie Sue
; APPLICANT: Thompson, Sheryl Ann
; TITLE OF INVENTION: GENE ENCODING CARBOXYPEPTIDASE OF
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5939305 of No. 5939305disk of No. 5939305th America, Inc.
; STREET: 405 Lexington Avenue, Suite 6400
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/967,149
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/608,452
; FILING DATE: 28-FEB-1996
; APPLICATION NUMBER: US 08/309,341
; FILING DATE: 20-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowney, Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 4247.000-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 867 0123
; TELEFAX: 212 867 0298
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2002 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: Aspergillus niger
; FEATURE:
; NAME/KEY: intron
; LOCATION: 349..411
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join (348..412)
; US-08-967-149-3

Alignment Scores:
Pred. No.: 1-9e-57 Length: 2002
Score: 623.00 Matches: 155
Percent Similarity: 48.25% Conservative: 79
Best Local Similarity: 31.96% Mismatches: 175
Query Match: 20.91% Indels: 76
DB: 2 Gaps: 18

US-09-712-338-2 (1-555) x US-08-967-149-3 (1-2002)

QY 45 AlaAsnAsnValThrIleArgTyrIysGluProGlyAlaGluGlyValCysGluThrThr 64
D 533 GCCTATGATCTCAGGGTCAAGAAGACCGATCCTAGCTCTCTTGGCATC-----GAC 593
QY 65 ProGlyValLysSerTyrSerGlyTyrValAspThrSerPro---GluSerHisThrPhe 83
D 584 CCTGGCGTAAAGCAGTACACCGGTTATCTCGATGACAGAGAACCGACAGCATCTGTTC 643
QY 84 PheTyrPhePheGluAlaArgHisAsnProGluThrAlaProIleThrLeuTyrLeuAsn 103
D 644 TACTGGTCTTTCGAGTCTCGCATGACCCGAGATGACCCCTGTTCTTGGCTGAAC 703
QY 104 GlyGlyProGlySerAspSerLeuIleGlyLeuPheGluGluLeuGlyProCysHisVal 123
D 704 GGTGGCCCTGGATGCTCTTCCCTCACCGGTCCTTTTCATGAGCTCGGCCCTAGCAGCATC 763
QY 124 AsnSerThrPheAspAspTyrIleAsnProHisSerTyrAsnGluValSerAsnLeuLeu 143
D 764 AACAAAGAGATCCAGCCGGTCTTACACGACTACGCTTGGAACTCCAAACGCGCTCGTATC 823
QY 144 PheLeuSerGlnProLeuGlyValGlyPheSerTyrSerAspThrValAspGlySerIle 163
D 824 TTCCTTGACCGCTGTCAACGTCGGTTACTCTTACAGCACTCTGCT-----871
QY 164 AsnProValThrGlyValValGluAsnSerSerPheAlaGlyValGlnGlyArgTyrPro 183
D 871 -----871

```

QY	184	ThrIleAspAlaThrLeuIleAspThrThrAsnLeuAlaAlaGluAlaAlaTrpGluIle	203
Db	872	---GTACGAGCACCGTGTCTGCTGGCAAGCAGGTCTATGCGCTTG	913
QY	204	LeuGlnGlyPheLeuSerGlyLeuProSerLeuAspSerArgValGlnSerLysAspPhe	223
Db	914	CTTACCTCTCTTCAACAATTCCTCC	961
QY	224	SerLeuTrpThrGluSerTyrGlyGlyHisTyrGlyProAlaPhePheAsnHisPheTyr	243
Db	962	CACATTGCGGTGAATCTATGCTGTCTACTATATCCCGCTTTGCTTGGAGATTG	1021
QY	244	GluGlnAsnGluArgIleAlaAsnGlySerValAsnGlyValGlnLeuAsnPheAsnSer	263
Db	1022	TCTCACAGAGAGCGC	1054
QY	264	LeuGlyIleIleAsnGlyIleIleAspGluAlaIleGlnAlaProTyrTyrProGluPhe	283
Db	1055	GTTCCTATTGGCAAGGTCTCACCAGCGTCTCACTCAGTACGAGTACTACCGTCCCATG	1114
QY	284	AlaValAsnAsnThrTyrGlyIleLysAlaValAsnGluThrValTyrAsnTyrMetLys	303
Db	1115	GCCTGTGTGACGGT--GGTTACCCAGCTGTCTGGACGAG--GGCTCTCTGCCAGGCC	1168
QY	304	PheAlaAsnAsnThrProAsnGlyCysGlnAspLeuIleSerThrCysLysGlnThrAsn	323
Db	1169	ATGGACAAACGCCCTTCCTCGC--TGCCAGCTCTATGATGTAGTCTTGCTATAGTTCCGAG	1225
QY	324	ArgThrAlaLeuAlaAspTyrAlaLeuCysAlaGluAlaThrAsnMetCysArgAspAsn	343
Db	1226	AGCGCT--TGGTTGTGTCCCGCTCCACTCTACTGTAAACAGCC	1270
QY	344	ValGluGlyProTyrTyrAlaPheAlaGlyArgGlyValTyrAspIleArgHisProTyr	363
Db	1271	CTCCTTGCCCTTACCAGCGC--ACCGACAGAACGCTACGATGTTCTGTGGTAAGTGC	1327
QY	364	AspAspPro-----ThrProSerTyrTyrAsnLysPheLeuAlaLys	378
Db	1328	GAGGATAGCTCAACCTCTGCTACTCGGCCATGGCTACGTCAGCGACTACTCTGACAAAG	1387
QY	379	AspSerValMetAspAlaIleGlyValAsnIleAsn--TyrThrGlnSerAsnAsnAsp	397
Db	1388	ACCGAGGTCAATGAGCTGTGGCGCTGAGGTCAACGGCTACGACTCGTGCACATTTGCAC	1447
QY	398	ValTyrTyrAlaPheGlnThrGlyAspPheValTyrPro-----AsnPheIleGlu	415
Db	1448	ATCAACCGCAACTTCTCTTCCACGCTGACTGGATGAAGCCCTACACCGTCTCGTTCGG	1507
QY	416	AspLeuGluGluIleLeuAlaLeuProValArgValSerLeuIleTyr--GlyAspAla	434
Db	1508	GGACTCTCTGGAG-----CAGATCTCTGTGC-----CTGATCTAGCTGGTGACGCC	1552
QY	435	AspTyrIleCysAsnTrpPheGlyGlyGlnAlaValSerLeuAlaAsnTyrSerGln	454
Db	1553	GATTTCACTGCAACTGGCTGGSCAACAGCGCTGCACTGAAGCCCTTGAGTGGCCGGA	1612
QY	455	AlaAlaGlnPheArgSerAlaGlyTyrThrProLeuLysVal-----Asn	469
Db	1613	CAGGCTGAATATGCTCCGCTAAGCTGGAGGACCTGGTGTGTCGAGAAATGAGCACAAAG	1672
QY	470	GlyValGluThrGlyGluThrArgGluTyrGlyAsnPheSerPheThrArgValTyrGlu	489
Db	1673	GCAGAAGATCGCGCAGGTCAAGTCCCATGGCACTTCACTTATGCGTCTCTATGCG	1732
QY	490	AlaGlyHisGluValProTyrTyrGlnProIleAlaSerLeuGlnLeuPheAsnArgThr	509
Db	1733	GGTGGCCACATGTCCTCCGATGGACCAACCGGATCGAGTCTTGAATCTTCAACCGCTGG	1792
QY	510	IlePheGly---Trp	513
Db	1793	TTGGGAGGTGAATGG	1807
RESULT 7			

US-08-309-341-1			
; Sequence 1, Application US/08309341			
; Patent No. 5594119			
; GENERAL INFORMATION:			
; APPLICANT: Yaver, Debbie Sue			
; APPLICANT: Thompson, Sheryl Ann			
; TITLE OF INVENTION: GENE ENCODING CARBOXYPEPTIDASE OF ASPERGILLUS NIGER			
; NUMBER OF SEQUENCES: 4			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: No. 55941190 No. 5594119disk of No. 5594119th America, Inc.			
; STREET: 405 Lexington Avenue, Suite 6400			
; CITY: New York			
; STATE: New York			
; COUNTRY: U.S.A.			
; ZIP: 10174-6401			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)			
; CURRENT APPLICATION DATA: US/08309,341			
; APPLICATION NUMBER: 16-SEP-1994			
; FILING DATE: 435			
; CLASSIFICATION: 435			
; ATTORNEY/AGENT INFORMATION:			
; NAME: Lowney, Karen A.			
; REGISTRATION NUMBER: 31,274			
; REFERENCE/DOCKET NUMBER: 4247.000-US			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: 212 867 0123			
; TELEFAX: 212 867 0298			
; INFORMATION FOR SEQ ID NO: 1:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 2068 base pairs			
; TYPE: nucleic acid			
; STRANDEDNESS: double			
; TOPOLOGY: linear			
; MOLECULE TYPE: Genomic DNA			
; ORIGINAL SOURCE:			
; ORGANISM: Aspergillus niger			
; FEATURE:			
; NAME/KEY: intron			
; LOCATION: 572..632			
; FEATURE:			
; NAME/KEY: CDS			
; LOCATION: join (571..633)			
US-08-309-341-1			
Alignment Scores:			
Pred. No.:	2,79e-54	Length:	2068
Score:	594.00	Matches:	155
Percent Similarity:	46.44%	Conservative:	80
Best Local Similarity:	30.63%	Mismatches:	174
Query Match:	19.94%	Indels:	98
DB:	1	Gaps:	18
US-09-712-338-2 (1-555) x US-08-309-341-1 (1-2068)			
QY	45	AlaAsnAsnValThrIleArgTyrLysGluProGlyAlaGluGlyValCysGluThrThr	64
Db	457	GCCTATGATCTCAGGCTCAAGAAGCCGATCTGGCTCTCTGGCATC-----GAC	507
QY	65	ProGlyValLysSerTyrSerGlyTyrValAspThrSerProGlu-----	79
Db	508	CCCGCGGTGAACAGTACACCGTTATCTCGATGACACGA-GAATGATAAGCATTTGTT	566
QY	80	-----SerHisThrPhePhe-----	84
Db	567	CTACGTAAGCACACCTTGGTTCACAGATCACGCTTTTATATGCTCTGGATATCTAACGCA	626
QY	85	-----TrpPheGluAlaArgHisAsnProGluThrAlaProIleThrLeuTrpLeu	102
Db	627	ACTTAGTGGTCTTCTCGAGTCTCGCAATGACCCCGAGATGATGATCCCGTCTGTGCGCTG	686


```
Query Match: 19.94% Indels: 98
DB: 1 Gaps: 18
US-09-712-338-2 (1-555) x US-08-608-267-1 (1-2068)
QY 45 AlaAsnValThrIleAspGlyValGluGlyValCysGluThr 64
DB 457 GCCTATGATCTCAGGCTCAAGAGACCGATCTCGCTCTCTTGGCATC-----GAC 507
QY 65 ProGlyValLysSerThrValAspThrSerProGlu-----79
DB 508 CCCGGCTGAAGCATACACCGGTTATCTCGATGACAAACA-GAATGATAAGCATTTGTT 566
QY 80 -----SerHisThrPhePhe-----84
DB 567 CTAGCTAAGCACACCTTGGTTCAAGATCACGCTTTTATATGCTCTGGATATCAACGCA 626
QY 85 -----TrpPheGluAlaArgHisAsnProGluThrAlaProIleThrLeuTyrPleu 102
DB 627 ACTTAGTGTTCTTCGAGTCTCGCAATGACCCCGAGATGATGATCGGTGTTCTCTGGCTG 686
QY 103 AsnGlyGlyProGlySerAspSerLeuIleGlyLeuPheGluGluLeuGlyProCysHis 122
DB 687 AACGGTGGCCCTGGGTCTCTCCCTACCGGTCTCTCATGAGCTTGGCCCTAGCAGC 746
QY 123 ValAsnSerThrPheAspAspTyrIleAsnProHisSerTrpAsnGluValSerAsnLeu 142
DB 747 ATCAACAAGAAAGATCCAGCCGGTCTACATGACTACGCTTGGAACTCCACGCGTCCGTG 806
QY 143 LeuPheLeuSerGlnProLeuGlyValGlyPheSerTyrSerAspThrValAspGlySer 162
DB 807 ATCTTCTCTGACCGCTCTCATGCTCGTTACTCTACAGTAATCTGCT-----857
QY 163 IleAsnProValThrGlyValGluAsnSerPheAlaGlyValGlnGlyArgTyr 182
DB 857 -----857
QY 183 ProThrIleAspAlaThrLeuIleAspThrThrAsnLeuAlaGluAlaThrPleu 202
DB 858 -----GTCAGCAGACGGTGGTGGCAAGACGCTCTATGCC 896
QY 203 IleLeuGlnGlyPheLeuSerGlyLeuProSerLeuAspSerArgValGlnSerLysAsp 222
DB 897 TTGCTTACCTCTTCTTCAAACTCC-----GAGTATGCTAAGCAGGAC 944
QY 223 PheSerLeuThrTrpGluSerTyrGlyHisTyrGlyProAlaPheAsnHisPhe 242
DB 945 TTCCACATTGCCGTGAATCTTATGCTGTGCTACTATATCCCGTCTTCTGCTCGGAGATC 1004
QY 243 TyrGluGlnAsnGluArgIleAlaAsnGlySerValAsnGlyValGlnLeuAsnPheAsn 262
DB 1005 CTGTCTCAACAAGAGCGC-----AACATCAACCTGCAG 1037
QY 263 SerLeuGlyIleIleAsnGlyIleIleAspGluAlaIleGlnAlaProTyrTrpGlu 282
DB 1038 TCGTCTCTATGGCAACGGTCTCACCGAGGATACACCCAGTACGAGTACTACGTCCTCC 1097
QY 283 PheAlaValAsnAsnThrTyrGlyIleLysAlaValAsnGluThrValTyrAsnTyrMet 302
DB 1098 ATGGCTCGGTGACGCGGCTTACCCAGCTGTCTTGAGCAGAGCTCTCCGCCATCTG 1157
QY 303 LysPheAlaAsnGlnMetProAsnGlyCysGlnAspLeuSerThrCysLysGlnThr 322
DB 1158 GAC-----AACGCTCTCTCCG-----TGCCAGTCTATGATTGAGTCTTGCTACAGTTCC 1208
QY 323 AsnArgThrAlaLeuAlaAspTyrAlaLeuCysAlaGluAlaThrAsnMetCysArgAsp 342
DB 1209 GAGAGCGCT-----TGGTTTGTGTCGCGGCTCTCTACTGTAACAAC 1253
QY 343 AsnValGluGlyProTyrTrpAlaPheAlaGlyValTyrAspIleArgHisPro 362
DB 1254 GCCTCTCTGCTTCCCGCTTACCGCGC-----ACTGGGCAAGACGCTATGATGTCCTGGTAAG 1310
```

```
QY 363 TyrAspAspPro-----ThrProSerTyrTyrAsnLysPheLeuAla 377
DB 1311 TGGGAGGATAGCTCTTAACCTTTGCTACTCGGTATGGCTTACGTCAGCGACTACCTGAAC 1370
QY 378 LysAspSerValMetAspAlaIleGlyValAsnIleAsn---TyrThrGlnSerAsnAsn 396
DB 1371 AAGCCCGAAGTCACTCAGGCTGTGGCGCTGAGGTCAACGGCTACGACTCGTGCACATTT 1430
QY 397 AspValTyrTrpAlaPheGlnThrGlyAspPheValTyrPro-----AsnPheIle 414
DB 1431 GACATCAACCGCAACTCTCTTCCACGGTGAGTGGATGAAGCCCTACCACCGCTCGTT 1490
QY 415 GluAspLeuGluGluIleLeuAlaLeuProValArgValSerLeuIleTyr---GlyAsp 433
DB 1491 CCGGAGCTCTCGGAG-----CAGATCCCTGTC-----TTGATCTATGCGCGTAT 1535
QY 434 AlaAspTyrIleCysAsnTrpPheGlyGlyAlaValSerLeuAlaAlaAsnTyrSer 453
DB 1536 GCTGATTTCTATTGCAACTGGCTGGGCAACAAGCGCTGGAGCTGAAGCCCTGGAGTGGCCC 1595
QY 454 GlnAlaAlaGlnPheArgSerAlaGlyTyrThrProLeuLysVal-----468
DB 1596 GGACAGCTGAAATATGCTCCGCTGAGCTGGAGGATCTGGTCAATTGTGCACATGAGCAC 1655
QY 469 AsnGlyValGluTyrGlyGluThrArgGluTyrGlyAsnPheSerPheThrArgValTyr 488
DB 1656 ACGGGCAAGAAGATTGGCCAGGTTAAGTCCCATGGCAACTTCACCTTCATCGCTCTCAT 1715
QY 489 GluAlaGlyHisGluValProTyrTyrGlnProIleAlaSerLeuGlnLeuPheAsnArg 508
DB 1716 GGTGTGGCCACATGGTCCCGATGGACGACCGCGAGTCTGAGTCTCGAGTCTTCAACCGC 1775
QY 509 ThrIlePheGly---Trp 513
DB 1776 TGGTTGGAGGTGATGG 1793
RESULT 9
US-08-608-452-1
; Sequence 1, Application US/08608452
; Patent No. 5693510
; GENERAL INFORMATION:
; APPLICANT: Yaver, Debbie Sue
; APPLICANT: Thompson, Sheryl Ann
; TITLE OF INVENTION: GENE ENCODING CARBOXYPEPTIDASE OF ASPERGILLUS NIGER
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 56935100 No. 5693510disk of No. 5693510th America, Inc.
; STREET: 405 Lexington Avenue, Suite 6400
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/608,452
; FILING DATE: 28-FEB-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/309,341
; FILING DATE: 20-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowney, Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 4247.000-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 867 0123
; TELEFAX: 212 867 0298
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
```

QY	263	SerLeuGlyIleIleAsnGlyValPheAlaSerGluAlaIleGlnAlaProTyrTyrProGlu	282
		: : : : : : : :	
Db	1038	TCCGTTCTCATGGCAACGGTCTCACCAGCGATACCCAGTAGTACTACCGTCCC	1097
		: : : : : : : :	
QY	283	PheAlaValAsnAsnThrTyrGlyIleLysAlaValAsnGluThrValTyrAsnTyrMet	302
		: : : : : : : :	
Db	1098	ATGGCTGCCTGACGCGGGTTACCCAGCTCTCTGGACGAGAGCTCTGCCAGTCCTCATG	1157
		: : : : : : : :	
QY	303	LysPheAlaAsnGlnMetProAsnGlyCysGlnAspLeulleSerThrCysLysGlnThr	322
		: : : : : : : :	
Db	1158	GAC-----AACGCTCTTCCTCGC---TGCGAGTCTATGATTGATGTCTGTACAGTTCC	1208
		: : : : : : : :	
QY	323	AsnArgThrAlaLeuAlaAspTyrAlaLeuCysAlaGluAlaThrAsnMetCysArgAsp	342
		: : : : : : : : :	
Db	1209	GAGAGCGCT-----TGGGTTTGTCCTCCGCGCTCCAFCTACTGTAAACAAC	1253
		: : : : : : : :	
QY	343	AsnValGluGlyProTyrTyrAlaPheAlaGlyArgGlyValTyrAspIleAArgHisPro	362
		:: : : : : : : : :	
Db	1254	GCCCTCTTGCCCTTACCAGCGC---ACTGGCGCAGAAGCTCTATGATGTCCTGGTAAG	1310
		: : : : : : : :	
QY	363	TyrAspAspPro-----ThrProProSerTyrTyrAsnLysPheLeuAla	377
		: : : : : : : : :	
Db	1311	TGCAGGATAGCTCTAACCTTTGCTACTCGGCTATGGCTACGTACGCGACTACCTGAAC	1370
		: : : : : : : :	
QY	378	LysAspSerValMetAspAlaIleGlyValAsnIleAsn--TyrThrGlnSerAsnAsn	396
		: : : : : : : :	
Db	1371	AAGCCCCAAGTCATCAGAGCTGTGTGGCGCTGAGGTCAAACGGCTACGACTCGTCACTTT	1430
		: : : : : : : :	
QY	397	AspValTyrTyrAlaPheGlnInThrGlyAspPheValTrpPro-----AsnPheIle	414
		: : : : : : : :	
Db	1431	GACATCAACCGCAACTTCTCTTCCAGGTGACTGGATGAAGCCCTACCACCGCTCGTIT	1490
		: : : : : : : :	
QY	415	GluAspLeuGluGluLeuAlaLeuProValArgValSerLeulleTyr---GlyAsp	433
		: : : : : : : :	
Db	1491	CCGGGATCTCTGAG-----CAGATCCCTGTC-----TTGATCTATGCGGTGAT	1535
		: : : : : : : :	
QY	434	AlaAspTyrIleCysAsnTrpPheGlyGlyGlnAlaValSerLeuAlaIleAsnTyrSer	453
		: : : : : : : :	
Db	1536	GCTGATTCATTTCGAATGGCTGGGCAACAGGCTTGAGCTGAAGCCCTGGAGTGGCCG	1595
		: : : : : : : :	
QY	454	GlnAlaAlaGlnPheArgSerAlaGlyTyrThrProLeuLysVal-----	468
		: : : : : : : :	
Db	1596	GGACAGCTGAATATGCTCCGCTGAGCTGGAGGATCTGGTCATTCGCACATGAGCAC	1655
		: : : : : : : :	
QY	469	AsnGlyValGluTyrGlyGluThrArgGluTyrGlyAsnPheSerPheThrArgValTyr	488
		: : : : : : : :	
Db	1656	ACGGGCAAGAAAGTGGCAGGTAAAGTCCCATGGCAACTTCACCTTCATGGGTCTCTAT	1715
		: : : : : : : :	
QY	489	GluAlaGlyHisGluValProTyrTyrGlnProIleAlaSerLeuGlnLeuPheAsnArg	508
		: : : : : : : :	
Db	1716	GTTGGTGGCCACATGTCCTCCGATGGACAGCCGAGTCGAGTCTCGAGTCTTCAACCGC	1775
		: : : : : : : :	
QY	509	ThrIlePheGly---Trp	513
		:: : : : : : : : :	
Db	1776	TGGTTGGAGGTGAATGG	1793
		: : : : : : : :	
RESULT 10			
US-08-608-224-1			
; Sequence 1, Application US/08608224			
; Patent No. 5705376			
; GENERAL INFORMATION:			
; APPLICANT: Yaver, Debbie Sue			
; APPLICANT: Thompson, Sheryl Ann			
; TITLE OF INVENTION: GENE ENCODING CARBOXYPEPTIDASE OF ASPERGILLUS NIGER			
; NUMBER OF SEQUENCES: 4			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: No. 5705376o No. 5705376disk of No. 5705376th America, Inc.			
; STREET: 405 Lexington Avenue, Suite 6400			
; CITY: New York			
; STATE: New York			
; COUNTRY: U.S.A.			
; ZIP: 10174-6401			
; COMPUTER READABLE FORM:			

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/608,224
; FILING DATE: 28-FEB-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/309,341
; FILING DATE: 20-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowney, Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 4247.000-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 867 0123
; TELEFAX: 212 867 0298
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2068 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Aspergillus niger
; FEATURE:
; NAME/KEY: intron
; LOCATION: 572..632
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join (571..633)
; US-08-608-224-1

```

Alignment Scores:

```

Pred. No.: 2,79e-54 Length: 2068
Score: 594.00 Matches: 155
Percent Similarity: 46.44% Conservativeness: 80
Best Local Similarity: 30.63% Mismatches: 174
Query Match: 19.94% Indels: 98
DB: 1 Gaps: 18

```

US-09-712-338-2 (1-555) x US-08-608-224-1 (1-2068)

```

QY 45 AlaAsnValThrIleArgTyrLysGluProGlyAlaGluGlyValCysGluThrThr 64
DB 457 GCCTATGATCTCAGGGTCAAGAAGACCGATCGTGGCTCTCTTGGCATC-----GAC 507
QY 65 ProGlyValLysSerTyrSerGlyTyrValAspThrSerProGlu-----79
DB 508 CCCGCGTGAAGCAGTACACCGGTTATCTCGATGACRACGA-GAATGATAAGCATTTGTT 566
QY 80 -----SerHisThrPhePhe-----84
DB 567 CTACGTAGACACACCTTGTTCAAGATCAGCGTTTTTATATGCTCTGGATATCTAACGCA 626
QY 85 -----TrpPheGluAlaArgHisAsnProGluThrAlaProIleThrLeuTrpLeu 102
DB 627 ACTTAGTGTTCTTCGAGTCTCGCAATACCCCGAGAATGATCCCGGTTGTTCTGTGGCTG 686
QY 103 AsnGlyGlyProGlySerAspSerLeuIleGlyLeuPheGluGluLeuGlyProCysHis 122
DB 687 AACGGTGGCCCTGGGTGCTCTCTCCCTACCGGTCTCTTCATGAGACTTGGCCCTAGCAGC 746
QY 123 ValAsnSerThrPheAspAspTyrIleAsnProHisSerTrpAsnGluValSerAsnLeu 142
DB 747 ATCAACAGAGAAGATCCAGCGGTCTACATGACTAGCTTGGAACTCCAAACCGTCCGCTG 806
QY 143 LeuPheLeuSerGlnProLeuGlyValGlyPheSerTyrSerAspThrValAspGlySer 162
DB 807 ATCTTCTTGACGACCGCTGCAATGCTGGTTACTCTACAGTAATCTGCT-----857

```

```

QY 163 IleAsnProValThrGlyValValGluAsnSerSerPheAlaGlyValGlnGlyArgTyr 182
DB 857 -----857
QY 183 ProThrIleAspAlaThrLeuIleAspThrThrAsnLeuAlaGluAlaAlaTrpGlu 202
DB 858 -----GTACGCGACACGCGTCTGCTGGCAAGGAGGAGCTCTATGCCC 896
QY 203 IleLeuGlnGlyPheLeuSerGlyLeuProSerLeuAspSerArgValGlnSerLysAsp 222
DB 897 TTGCTTACCCTCTCTTCAACAATTCCTCC-----GAGTATGTAAGCAGAC 944
QY 223 PheSerLeuTrpThrGluSerTyrGlyGlyHisTyrGlyProAlaPheAsnHisPhe 242
DB 945 TTCCACATTCGCGGTGAATCTTATGCTGGTCACTATATCCCGCTCTTCGCTTCGAGATC 1004
QY 243 TyrGluGlnAsnGluArgIleAlaAsnGlySerValAsnGlyValGlnLeuAsnPheAsn 262
DB 1005 CTGTCTCACAGAAGAGCGC-----AACATCAACCTCGAG 1037
QY 263 SerLeuGlyIleIleAsnGlyIleIleAspGluAlaIleGlnAlaProTyrTrpProGlu 282
DB 1038 TCGGTTCTTATGGCAACGGTCTCACCGAGGATACACCGAGTACGAGTACTACCGTCCC 1097
QY 283 PheAlaValAsnAsnThrTyrGlyIleLysAlaValAsnGluThrValTyrAsnTyrMet 302
DB 1098 ATGSCCTGCGGTGACGGCGGTACCAGCTGTCTTGGACGAGAGCTCTGCGAGTCCATG 1157
QY 303 LysPheAlaAsnGlnMetProAsnGlyCysGlnAspLeuIleSerThrCysLysGlnThr 322
DB 1158 GAC-----AACGCTCTTCTCTCGC---TGGCAGCTTATGATGAGTCTTGTACAGTTC 1208
QY 323 AsnArgThrAlaLeuAlaAspTyrAlaLeuCysAlaGluAlaThrAsnMetCysArgAsp 342
DB 1209 GAGAGCGCT-----TGGGTTGTCTCCGCGCTCCACTACTGTATAACAAC 1253
QY 343 AsnValGluGlyProTyrTyrAlaPheAlaGlyArgGlyValTyrAspIleArgHisPro 362
DB 1254 GCCCTCTTGGCCCTTACCAGCGC---ACTGGGAGAAACGCTCTATGATGCTCGTGGTAA 1310
QY 363 TyrAspAspPro-----ThrProSerTyrTyrAsnLysPheLeuAla 377
DB 1311 TGCAGGATAGCTCTAACTTGTCTACTCGGTATGGGTACGTACGAGGACTACTCTGAAC 1370
QY 378 LysAspSerValMetAspAlaIleGlyValAsnIleAsn-----TyrThrGlnSerAsnAsn 396
DB 1371 AAGCCCGAGTATCGAGGCTGTGGCGCTCAGGTCAACGCTACGACTCTCGCAACTTT 1430
QY 397 AspValTyrTyrAlaPheGlnGlnThrGlyAspPheValTrpPro-----AsnPheIle 414
DB 1431 GACATCAACCGCAACTTCTCTCCACGCTGACTGGATGAAGCGCTACACCGCTCGT 1490
QY 415 GluAspLeuGluIleLeuAlaLeuProValArgValSerLeuIleTyr---GlyAsp 433
DB 1491 CCGGACTCTCGAG-----CAGATCCCTGTC-----TTGATCTATCCCGGTAT 1535
QY 434 AlaAspTyrIleCysAsnTrpPheGlyGlyGlnAlaValSerLeuAlaAsnTrpSer 453
DB 1536 GCTGATTTCAATTGCACTGCTGGCAACAGGCTCGACTGAAGCCCTGGAGTGGCCC 1595
QY 454 GlnAlaAlaGlnPheArgSerAlaGlyTyrThrProLeuLysVal-----468
DB 1596 GGACAGGCTGAATATGCTCGGCTGAGCTGGAGGATCTGGTCATTGTGCAANTGAGCAC 1655
QY 469 AsnGlyValGluTyrGlyGluThrArgGluTyrGlyAsnPheSerPheThrArgValTyr 488
DB 1656 ACGGCAAGAGATTGGCCAGGTTAAGTCCCATGCGCACTTCACTTCCATGCTCTCTAT 1715
QY 489 GluAlaGlyHisGluValProTyrTyrGlnProIleAlaSerLeuGlnLeuPheAsnArg 508
DB 1716 GGTGGTGGCCATGTTCCCGATGACACCGCGAGTGGAGTCTCGAGTCTCTCAACCGC 1775
QY 509 ThrIlePheGly---Trp 513

```

```

Db      1776 TGGTTGGAGGTGAATGG 1793
RESULT 11
US-08-967-149-1
; Sequence 1, Application US/08967149
; Patent No. 5939305
; GENERAL INFORMATION:
; APPLICANT: Yaver, Debbie Sue
; APPLICANT: Thompson, Sheryl Ann
; TITLE OF INVENTION: GENE ENCODING CARBOXYPEPTIDASE OF
; TITLE OF INVENTION: ASPERGILLUS NIGER
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESS: No. 59393050 No. 5939305disk of No. 5939305th America, Inc.
; STREET: 405 Lexington Avenue, Suite 6400
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/967,149
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/608,452
; FILING DATE: 28-FEB-1996
; APPLICATION NUMBER: US 08/309,341
; FILING DATE: 20-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowney, Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 4247.000-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 867 0123
; TELEFAX: 212 867 0298
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2068 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Aspergillus niger
; FEATURE:
; NAME/KEY: intron
; LOCATION: 572..632
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join (571..633)
; US-08-967-149-1
Alignment Scores:
Pred. No.: 2,79e-54 Length: 2068
Score: 594.00 Matches: 155
Percent Similarity: 46.44% Conservative: 80
Best Local Similarity: 30.63% Mismatches: 174
Query Match: 19.94% Indels: 98
DB: 2 Gaps: 18
US-09-712-338-2 (1-555) x US-08-967-149-1 (1-2068)
QY      45 AlaAsnValThrIleArgTyrLysGluProGlyAlaGluGlyValCysGluThrThr 64
Db      457 GCCTATGATCTCGAGTCAAGAACCGATCCCTGGCTCTCTGGCATC-----GAC 507
QY      65 ProGlyValLysSerTyrSerGlyTyrValValAspThrSerProGlu----- 79

```

1431 GACATCAACCGCACTTCTTCCACGGTCACTGGATGAGCCCTACCACCGCTCGTT 1490
415 GluAspLeuGluLeuLeuAlaLeuProValArgValSerLeuLeuTyr---GlyAsp 433
1491 CCGGACTCTCTGGAG-----CAGATCCCTGTC-----TTGATCTATGCCGGTGAT 1535
434 AlaAspTyrIleCysAsnTrpPheGlyGlyGlnAlaValSerLeuAlaAlaAsnTyrSer 453
1536 GCTGATTTCATTTCAATTCGCTGGCAACAGCGCTGGAGCTGAGCCCTGGATGGGCC 1595
454 GlnAlaAlaGlnPheArgSerAlaGlyTyrThrProLeuLysVal----- 468
1596 GCACAGCTGATATGCTCCGCTCAGCTGGAGGATCTGGTCATGTGCACATGAGCAC 1655
469 AsnGlyValGluTyrGlyGluThrArgGluTyrGlyAsnPheSerPheThrArgValTyr 488
1656 ACGGCAAGAAAGATTGGCCAGGTAAAGTCCCATGCCAACTTCACCTTCATCGCTCTAT 1715
489 GluAlaGlyHisGluValProTyrTyrGlnProIleAlaSerLeuGlnLeuPheAsnArg 508
1716 GGTGTGGCCACATGCTCCCATGACGACCGCCGAGTGGAGTCTCGAGTTCTTCAACGC 1775
509 ThrIlePheGly---Tip 513
1776 TGGTGGGAGTGAATGG 1793

RESULT 12
US-09-640-305-3
; Sequence 3, Application US/09640305
; Patent No. R837447
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; Fournier, Alain
; Yeh, Patrice
; TITLE OF INVENTION: MODIFIED KLUYVEROMYCES YEASTS, THEIR
; PREPARATION AND USE
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Rd. 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19002
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/640,305
; FILING DATE: 16-Aug-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/360,673
; FILING DATE: 06-FEB-1995
; APPLICATION NUMBER: WO PCT/FR93/00623
; FILING DATE: 23-JUN-1993
; APPLICATION NUMBER: FR 92/07785
; FILING DATE: 25-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, Julie K.
; REGISTRATION NUMBER: 38,619
; REFERENCE/DOCKET NUMBER: ST92040-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610)454-3839
; TELEFAX: (610)454-3808
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2503 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA

ORIGINAL SOURCE:
ORGANISM: Kluyveromyces lactis
FEATURE:
NAME/KEY: CDS
LOCATION: 387..1862
OTHER INFORMATION: /product= "K. lactis protease C
gene"
/gene= "Kl.PRC1"
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-640-305-3

Alignment Scores:
Pred. No.: 2,2e-51 Length: 2503
Score: 568.50 Matches: 151
Percent Similarity: 44.65% Conservative: 66
Best Local Similarity: 31.07% Mismatches: 184
Query Match: 19.08% Indels: 85
DB: 1 Gaps: 14

US-09-712-338-2 (1-555) x US-09-640-305-3 (1-2503)
Qy 45 AlaAsnAsnValThrIleArgTyrLysGluProGlyAlaGluGlyValCysGluThrThr 64
Db 570 GCTTATTCAATTGAGATTAACCTTTGGATCCCAATCTCTTGGCTT----- 617
Qy 65 ProGlyValLysSerTyrSerGlyTyrValAspThrSerProGluSerHisThrPhePhe 84
Db 618 GATACCGTGAACCAATGGTCGGATATTAGATTACCAGGACTCAAAACACTCTCTTTAT 677
Qy 85 TrpPhePheGluAlaArgHisAsnProGluThrAlaProIleThrLeuTrpLeuAsnGly 104
Db 678 TGGTTTTTGGAGTCTAGAAATGACCCAGAGAATGACCCAGTATGATATGTTAAACGGT 737
Qy 105 GlyProGlySerAspSerLeuIleGlyLeuPheGluGluGlyProCysHisValAsn 124
Db 738 GGTCTGGCTGCTTCTCTCTCGTGGTCTTTTCTTGAATTGGGACCTTCTCTATAGA 797
Qy 125 SerThrPheAspTyrIleAsnProHisSerTrpAsnGluValSerAsnLeuLeuPhe 144
Db 798 GCTGATTTGAAACCACTTATAACCCCTACTCTTGGATTCCACGCTTCTGTGATATTC 857
Qy 145 LeuSerGlnProLeuGlyValGlyPheSerTyrSerAspThrValAspGlySerIleAsn 164
Db 858 CTAGATCAGCCTGTTGGTGGTCTCTCATACCGGTGAC----- 896
Qy 165 ProValThrGlyValValGluAsnSerSerPheAlaGlyValGlnGlyArgTyrProThr 184
Db 896 ----- 896
Qy 185 IleAspAlaThrLeuIleAspThrThrAsnLeuAlaAlaGluAlaAlaTrpGluLeu 204
Db 897 -----TCTAAGTGTCTACTACAGATGACGCTGCCAAGACGCTTACATATCTTAT 947
Qy 205 GlnGlyPheLeuSerGlyLeuProSerLeuAspSerArgValGlnSerLysAspPheSer 224
Db 948 GATTTGTCTTTGAAAGATTCCCTCATTTG-----AGAAATAACCAATTCAT 995
Qy 225 LeuTrpThrGluSerTyrGlyHisTyrGlyProAlaPhePheAsnHisPheTyrGlu 244
Db 996 ATCTCCGGTGAATCATACCGCGCTCATTTATTACCAAGATTGCTCAT----- 1043
Qy 245 GlnAsnGluArgIleAlaAsnGlySerValAsnGlyValGlnLeuAsnPheAsnSerLeu 264
Db 1044 -----GAGATTGCTGTAGTGCATGCTGAGGATTCCTCTCAATCTATCGTCAGTA 1094
Qy 265 GlyIleLeuAsnGlyIleLeuAspGluAlaIleGlnAlaProTyrTyrProGluPheAla 284
Db 1095 TTAATTGGAATGGATTTACTGACCCACTGACTCAATACCAATATATACGAGCGGATGGCC 1154
Qy 285 ValAsnAsnThrTyrGlyIleLysAlaValAsnGluThrValTyrAsnTyrMetLysPhe 304
Db 1155 TGTGTGAAGGT---GGTTATCCAGCGGTGTTGGAA----- 1187


```

QY 305 AlaAsnGlnMetProAsnGlyCysGlnAsp----- 314
Db 1188 -----COGGAAGATTGCTTAGATATGAATAGGAATCTACCTCTATGCCTATCG 1235
QY 315 LeuIleSerThrCysLysGlnThrAsnArgThrAlaLeuAlaAspTyrAlaLeuCysAla 334
Db 1236 CTTGTGGACCGCTGTTACAAGTCCCAT-----TCTGTTTTCTCTGTGTG 1280
QY 335 GluAlaThrAsnMetCysArgAspAsnValGluGlyProTyrTyrAlaPheAlaGlyArg 354
Db 1281 TTGGCTGACCGTTATTGTGAACAACAGATTACTGGG---GTTTATGAGAAATCAGGTAGG 1337
QY 355 GlyValTyrAspIleArg-----HisProTyrAspAspPro-----Thr 367
Db 1338 AACCTTACGATATTAGATCTAAGTGTGAAGCAGAGATGATCCGGTGCCTGTTATCAG 1397
QY 368 ProProSerTyrTyrAsnLysPheLeuAlaLysAspSerValMetAspAlaIleGlyVal 387
Db 1398 GAAGAAATTTATATCTCTGATTACTTGAATCAGGAGGAAGTTCAAAGAGCTTTAGGACT 1457
QY 388 AsnIleAsnTyrThrGlnSer---AsnAspAspValTyrTyrAlaPheGlnGlnThrGly 406
Db 1458 GATGTGAGTCTTCCAGGTTGATCGCTCGGATCGGTATCGGTTTCGCAITCAGCTGGC 1517
QY 407 AspPheValTrpProAsnPheIleGluAspLeuGluGluIleLeuAlaLeuProValArg 426
Db 1518 GATGGACCGAGCCCA---TTCCACCACTAGCTGCAGAACTTCTTGATCAAGATATCAAT 1574
QY 427 ValSerLeuIleTyrGlyAspAlaAspTyrIleCysAsnTrpPheGlyGlyGlnAlaVal 446
Db 1575 GTCTTGATATANGCAGCGGATAAGGATTATATTGTAATTGGCTAGGAATCTCGCTGG 1634
QY 447 SerLeuAlaAlaAsnTyrSerGlnAlaAlaGlnPheArgSerAlaGlyTyrTrpProLeu 466
Db 1635 ACTGAAAAATGGAATGGAGGTATACGAAGAGTATATAAAAAACAAGTTTTCAGAACTGG 1694
QY 467 LysValAsnGlyValGlu-----TyrGlyGluThrArgGluTyrGlyAsnPheSerPhe 484
Db 1695 AAGAGTGAAGAAACAGACAGACCATTTGGGAAACCAATCTTATGGCCGCTAACTTAC 1754
QY 485 ThrArgValTyrGluAlaGlyHisGluValProTyrTyrGlnProIleAlaSerLeuGln 504
Db 1755 TTGAGAAATCTATGATGCTGACACATGTTCTCTCAGACCAACCTGAAAAATTCATTACAA 1814
QY 505 LeuPheAsnArgThrIle 510
Db 1815 ATGGTGAATTCATGAAT 1832

RESULT 13
US-08-360-673-3
; Sequence 3, Application US/08360673
; Patent No. 5679544
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; APPLICANT: Fournier, Alain
; APPLICANT: Yeh, Patrice
; TITLE OF INVENTION: MODIFIED KLUYVEROMYCES YEASTS, THEIR
; PREPARATION AND USE
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Rd. 3043
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19002
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/360,673

```

```

; FILING DATE: 435
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/FR93/00623
; FILING DATE: 23-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 92/07785
; FILING DATE: 25-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, Julie K.
; REGISTRATION NUMBER: 38,619
; REFERENCE/DOCKET NUMBER: ST92040-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610)454-3839
; TELEFAX: (610)454-3808
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2503 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: Kluyveromyces lactis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 387..1862
; OTHER INFORMATION: /product= "K. lactis protease C
; OTHER INFORMATION: gene"
; OTHER INFORMATION: /gene= "Kl.PRC1"
; US-08-360-673-3

Alignment Scores:
Pred. No.: 2,2e-51 Length: 2503
Score: 568.50 Matches: 151
Percent Similarity: 44.65% Conservative: 66
Best Local Similarity: 31.07% Mismatches: 184
Query Match: 19.08% Indels: 85
DB: 1 Gaps: 14

US-09-712-338-2 (1-555) x US-08-360-673-3 (1-2503)
QY 45 AlaAsnValThrIleArgTyrLysGluProGlyAlaGluGlyValCysGluThrThr 64
Db 570 GCTTATTCAATGAGAAATTAACCTTTGGATCCCAATCTCTTGGCGTT----- 617
QY 65 ProGlyValLysSerTyrSerGlyTyrValAspThrSerProGluSerHisThrPhePhe 84
Db 618 GATACCGTGAACAACATGGTCCGGATATTAGATTACCAGAGACTCAAAACACTCTTTTAT 677
QY 85 TrpPheGluAlaArgHisAsnProGluThrAlaProIleThrIleuTrpLeuAsnGly 104
Db 678 TGGTTTTTTGAGCTAGAAATGACCCAGAGAATGACCCAGATGATATGTTAAACGGT 737
QY 105 GlyProGlySerAspSerLeuIleGlyLeuPheGluGluGlyProCysHisValAsn 124
Db 738 GGTCTGGCTGTCTCTTTCGTCTTCTTTGAAATGGGACCTTCTTCTATAGGA 797
QY 125 SerThrPheAspTyrIleAsnProHisSerTrpAsnGluValSerAsnLeuLeuPhe 144
Db 798 GCTGATTTGAAACCCATTTATACCCCTACTCTTGGAAATCCACGCTTCTGTGATTC 857
QY 145 LeuSerGlnProLeuGlyValGlyPheSerTyrSerAspThrValAspGlySerIleAsn 164
Db 858 CTAGATCAGCCTGTGGTGTGGTCTCTCATACGGTGAC----- 896
QY 165 ProValThrGlyValValGluAsnSerSerPheAlaGlyValGlnGlyArgTyrProThr 184
Db 896 ----- 896
QY 185 IleAspAlaThrLeuIleAspThrThrAsnLeuAlaAlaGluAlaAlaTrpGluIleLeu 204
Db 897 -----TCTAAAGTGTCTACTACAGATGCGCTGCCAAGACGTTTACATATCTTAA 947

```

QY 205 GlnGlyPheLeuSerGlyLeuProSerLeuAspSerArgValGlnSerLysAspPheSer 224
 Db 948 GATTGTGTTCTTGAAGATTCCCTCATTTG-----AGAAATAACGATTTCAT 995
 QY 225 LeuTrpThrGluSerTyrglyGlyHisTyrglyProAlaPhePheAsnHisPheTyrgly 244
 Db 996 ATCTCCGGTGAATCATACGCGGTCATTATTTACCAAGATTGCTCAT----- 1043
 QY 245 GlnAsnGluArgIleAlaAsnGlySerValAsnGlyValGlnLeuAsnPheAsnSerLeu 264
 Db 1044 -----GAGATTGCTGAGTGCATCTGAGGATTCCTTCATCATCTATCTCAGTA 1094
 QY 265 GlyIleIleAsnGlyIleIleAspGluAlaIleGlnAlaProTyrglyProGluPheAla 284
 Db 1095 TTAATTGGAATGATTACTAGCCACTGACTCAATACCAATATTACGAGCCGATGCC 1154
 QY 285 ValAsnAsnThrTyrglyIleLeuAlaValAsnGluThrValTyrglyMetLysPhe 304
 Db 1155 TGCTGGAAGT---GGTTATCCAGCGGTGTGGA----- 1187
 QY 305 AlaAsnGlnMetProAsnGlyCysGlnAsp----- 1187
 Db 1188 -----CCGGAAGATTGCTTAGATATGAATAGGAATCTACCTCTATGCTATCG 1235
 QY 315 LeuIleSerThrCysLysGlnThrAsnArgThrAlaLeuAlaAspTyrglyAlaLeuCysAla 334
 Db 1236 CTGTGGACCGCTTACAGTCCCAT-----TCTGTTTCTCTGTGTG 1280
 QY 335 GluAlaThrAsnMetCysArgAspAsnValGluGlyProTyrglyAlaPheAlaGlyArg 354
 Db 1281 TTGCTGACCGTTATTGTGAACAACAGATTACTGG---GTTATGAGAATCAGGTAGG 1337
 QY 355 GlyValTyrglyAspIleArg-----HisProTyrglyAspPro-----Thr 367
 Db 1338 AACCTTACGATATTAGATCTAAGTGAAGCAGAGGATGATCCGGTGCCTGTTATCAG 1397
 QY 368 ProSerSerTyrglyAsnLysPheLeuAlaLysAspSerValMetAspAlaIleGlyVal 387
 Db 1398 GAAGAAATTTATCTCTGATTACTTGAATCAGAGGAGGTTCAAGAGCTTTAGGGA 1457
 QY 388 AsnIleAsnTyrglyGlnSer---AsnAsnAspValTyrglyAlaPheGlnGlnThrGly 406
 Db 1458 GATGTGAGTCTTTCCAGAGTTGTAGTCGGATGTCGGTTCGGATTCAGTGGC 1517
 QY 407 AspPheValTrpProAsnPheIleGluAspLeuGluIleLeuAlaLeuProValArg 426
 Db 1518 GATGACCGAGCCCA---TTCACCCAGTACGTCGAGAACTTCTTGATCAAGATATCAAT 1574
 QY 427 ValSerLeuIleTyrglyAspAlaAspTyrglyCysAsnTrpPheGlyGlyGlnAlaVal 446
 Db 1575 GTCTGTATATGACGCGTAAGGATTAATTTGTAATGGCTAGGAAATCTCGCTGG 1634
 QY 447 SerLeuAlaAlaAsnTyrglySerGlnAlaAlaGlnPheArgSerAlaGlyTyrglyProLeu 466
 Db 1635 ACTGAAATTTGATGAGGTATTAACGAGAGTATAAAACAAAGTTTTCGAACTTGG 1694
 QY 467 LysValAsnGlyValGlu-----TyrglyGluThrArgGluTyrglyAsnPheSerPhe 484
 Db 1695 AAGAGTGAAGAACAGATGAGACCAATTCGCGAAACCAATCTTATGGCCCGCTAATTC 1754
 QY 485 ThrArgValTyrglyAlaGlyHisGluValProTyrglyGlnProIleAlaSerLeuGln 504
 Db 1755 TTGAAATCTATGATGCTGGACATATGTTCTTCACGCCAACCTGAAATTCATTA 1814
 QY 505 LeuPheAsnArgThrIle 510
 Db 1815 ATGGTGAATTCATGGATT 1832

RESULT 14

US-08-899-324-32

; Sequence 32, Application US/08899324

; Patent No. 5945329

GENERAL INFORMATION:
 APPLICANT: Breddam, Klaus
 APPLICANT: Keilland-Brandt, Morten
 APPLICANT: Mortensen, Uffe
 APPLICANT: Olesen, Kjeld
 APPLICANT: Stennicke, Henning
 APPLICANT: Wagner, Fred
 TITLE OF INVENTION: CUSTOMIZED PROTEASES
 NUMBER OF SEQUENCES: 33
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt
 STREET: 3100 No. 5945329west Center, 90 S. 7th Street
 CITY: Minneapolis
 STATE: MN
 COUNTRY: U.S.A.
 ZIP: 55402
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq Version 1.5
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/899,324
 FILING DATE: 23-JUL-1997
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/329,892
 FILING DATE: 27-OCT-1994
 APPLICATION NUMBER: 08/144,704
 FILING DATE: 28-OCT-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Kettlebeiger, Denise M
 REGISTRATION NUMBER: 33,924
 REFERENCE/DOCKET NUMBER: 8648.44USC1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 612/332-5300
 TELEFAX: 612/332-9081
 TELEX:
 INFORMATION FOR SEQ ID NO: 32:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2632 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: Genomic DNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 FRAGMENT TYPE:
 ORIGINAL SOURCE:
 FEATURE:
 NAME/KEY: Coding Sequence
 LOCATION: 696..2291
 OTHER INFORMATION:
 US-08-899-324-32

Alignment Scores:
 Pred. No.: 1,81e-46 Length: 2632
 Score: 523.50 Matches: 147
 Percent Similarity: 43.68% Conservative: 74
 Best Local Similarity: 29.05% Mismatches: 200
 Query Match: 17.57% Indels: 85
 DB: 2 Gaps: 19

US-09-712-338-2 (1-555) x US-08-899-324-32 (1-2632)

QY 33 ProLysAsnProThrGlyValIleThr-----LeuThrThrAlaAsnAsn 47
 Db 939 CCAAAATTCCTGAAGCAATCAAAACGAAGAAAGACTGGGACTTTGTGTCAGAAATGAC 998
 QY 48 ValThrIleArgTyrgly-----LysGluProGlyAlaGluGlyVal 60
 Db 999 GCAATTTGAAACTATCAGCTTCGTGTCACAAAGATTAGGACCTTAAATCTGGGCATT 1058

```

QY 61 CysGluThrThrProGlyValLysSerTyrSerGlyTyrValAspThrSerProGlu--- 79
Db 1059 -----GACCACAAATGTCACACAGTACACGGGTACTTGGATGTGGAGACGAGAC 1109
QY 80 SerHisThrPhePheThrPhePheGluAlaArgHisAsnProGluThrAlaProIleThr 99
Db 1110 AAGCATTTCTCTTTTGGACTTTTGAAGTAGAAGACATCTGCAAGAGGATCCGGTCATC 1169
QY 100 LeuTrpLeuAsnGlyGlyProGlySerAspSerLeuIleGlyLeuPheGluGluLeuGly 119
Db 1170 CTITGGTGTGAACGGGGTCCAGGTTGTCTTCACTAACCGGGCTGTCTTTTGAATAGGA 1229
QY 120 ProCysHisValAsnSerThrPheAspAspTyrIleAsnProHisSerTyrAsnGluVal 139
Db 1230 CCCTCATCCATTGGACCTGATTGAAACCCATCGGAAACCTTACTCTTGGAAACAGCAAT 1289
QY 140 SerAsnLeuLeuPheLeuSerGlnProLeuGlyValGlyPheSerTyrSerAspThrVal 159
Db 1290 GCCACCGTGATCTCTCTTGACACAGCTGTCAACGTTGGTTCGTATTCC----- 1340
QY 160 AspGlySerIleAsnProValThrGlyValValGluAsnSerSerPheAlaGlyValGln 179
Db 1341 -----GGGTCCCTCA 1349
QY 180 GlyArgTyrProThrIleAspAlaThrLeuIleAspThrThrAsnLeuAlaAlaGluAla 199
Db 1350 GGT-----GTTTCCACACTGTGCGCGCTGTGATAGAT 1382
QY 200 AlaTrpGluIleLeuGlnGlyPheLeuSerGlyLeuProSerLeuAspSerArgValGln 219
Db 1383 GTCTATAACTTCTTGGAGTTGTTCTTCGATCAGTTCCTCAATAGCGTCAACAAAGGGCAA 1442
QY 220 SerIlyAspPheSerLeuTrpThrGluSerTyrGlyHisTyrGlyProAlaPhePhe 239
Db 1443 -----GATTTCCATCTCGTGGGAATCTACCGCGCATTACATCCCTGTTTGGC 1496
QY 240 AsnHisPheTyrGluGlnAsnGluArgIleAlaAsnGlySerValAsnGlyValGlnLeu 259
Db 1497 TCTGAAATTTGCTCACAAGGACAG----- 1523
QY 260 AsnPheAsnSerLeuGlyIlelle-----AsnGlyIlelleAspGluAlalleGlnAla 277
Db 1524 AACTTCAACTTAACCTCCGCTGTGCGGAATGGCGTCACTGACCCATTGACTCAGTAT 1583
QY 278 ProTyrTyrProGluPheAlaValAsnAsnThrTyrGlyIleIlysalValAsnGluThr 297
Db 1584 AACTATTACAAACCAATGGCCTGTGTGAAGGT---GGCGAACCATCTGTT----- 1631
QY 298 ValTyrAsnTyrMetLysPheAlaAsnGlnMetProAsnGlyCysGlnAspLeuIleSer 317
Db 1632 -----TTGCCCTCGGAGGATGCTCTGCTATGGAAGACTCTTTGGAA 1673
QY 318 ThrCysGlnThrAsnArgThrAlaLeuAlaAspTyrAlaLeu-----CysAlaGlu 335
Db 1674 CGTTGTTGGGCTGTGATCGAGTCGTCGATGACTCGCAATCGGTCTGCTGTGTTCCA 1733
QY 336 AlaThrAsnMetCysArgAspAsnValGluGlyProTyrTyrAlaPheAlaGlyArgGly 355
Db 1734 GCTACCATTTATGTAATACAGCCCAATTTGGCTTCTTACCAAGT---ACCGGCAGAAAC 1790
QY 356 ValTyrAspIleArgHisProTyrAspAsp-----ProThrProProSer 370
Db 1791 GTTTACGATACAGGAAGGATTGTGAAGTGGCAATTTGTGCTACCAAGTTACAGAT 1850
QY 371 TyrTyrAsnLysPheLeuAlaLysAspSerValMetAspAlalleGlyValAsnIleAsn 390
Db 1851 ---ATCGACGACTACTTAAACACGAGTACGTCAAGAGAGCTGTCGTGGGAGGTTGAC 1907
QY 391 TyrThrGlnSer---AsnAsnAspValTyrTyrAlaPheGlnGlnThrGlyAspPheVal 409
Db 1908 CACTAGAAATCCTGTAATCTCGATCAACAGAAATTTCTGTTTTCGGGTGATTTGGATG 1967
QY 410 TrpProAsnPheIleGluAspLeuGluIleLeuAlaLeuProValArgValSerLeu 429

```

```

Db 1968 AAGCCT---TACCACACCGCCGTAACAGATCTTTTGAATCAAGACCTTACCATTCTCGTA 2024
QY 430 IleTyrGlyAspAlaAspTyrIleCysAsnTyrPheGlyGlyGlnAlaValSerLeuAla 449
Db 2025 TATGACGGCGATAAAGATTTCATCTGTAACCTGGTGGTAAATAGGGGTGACCGATGTC 2084
QY 450 AlaAsnTyrSerGlnAlaAlaGlnPheArgSer-----AlaGlyTyrThrProLeu 466
Db 2085 TTGCCATGGAAGTACGACGAGAAATTTGCAAGCAAAAGTAGCTAACTGACTGCTTCT 2144
QY 467 LysValAsnGlyValGluTyrGlyGluThrArgGluTyrGlyAsnPheSerPheThrArg 486
Db 2145 ATCACCAGCAGAGTCGCT---GGTGAAGTCAATCTTACAGCACATTCACCTATTGAGA 2201
QY 487 ValTyrGluAlaGlyHisGluValProTyrTyrGlnProIleAlaSerLeuGlnLeuPhe 506
Db 2202 GTCTTCAATGGTGGCCACATGGTTCATTTGACGTCCCTGAAACCGCTTAAGTATGGTT 2261
QY 507 AsnArgThrIlePheGly 512
Db 2262 AACGAATGGATCCACGGT 2279
RESULT 15
US-08-329-892B-32
; Sequence 32, Application US/08329892B
; Patent No. 6187579
; GENERAL INFORMATION:
; APPLICANT: Breddam, Klaus
; APPLICANT: Keilland-Brandt, Morten
; APPLICANT: Mortensen, Uffe
; APPLICANT: Olesen, Kjeld
; APPLICANT: Stennicke, Henning
; APPLICANT: Wagner, Fred
; TITLE OF INVENTION: CUSTOMIZED PROTEASE
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt
; STREET: 3100 No. 618/579west Center, 90 S. 7th Street
; CITY: Minneapolis
; STATE: MN
; COUNTRY: U.S.A.
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/329,892B
; FILING DATE: 27-OCT-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/144,704
; FILING DATE: 28-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Kettleberger, Denise M
; REGISTRATION NUMBER:
; REFERENCE/DOCKET NUMBER: 8648.44US01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612/332-5300
; TELEFAX: 612/332-9081
; TELEX:
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2632 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE:

```

ORIGINAL SOURCE:

; ; NAME/KEY: Coding Sequence

; ; FEATURE: ; ; LOCATION: 696...2291

; ; OTHER INFORMATION:

US-08-329-892B-32

Alignment Scores:

Pred. No.:	1.81e-46	Length:	2632
Score:	523.50	Matches:	147
Percent Similarity:	43.68%	Conservative:	74
Best Local Similarity:	29.05%	Mismatches:	200
Query Match:	17.57%	Indels:	85
DB:	3	Gaps:	19

US-09-712-338-2 (1-555) x US-08-329-892B-32 (1-2632)

```

QY 33 ProLysAsnProThrGlyValLysThr-----LeuThrThrAlaAsnAsn 47
Db 939 CCAAAATTCCTGAAGCAATCAAAACGAAGAAAGACTGGGACTTTGTGTCAGAGATGAC 998
QY 48 ValThrIleArgTyr-----LysGluProGlyAlaGluGlyVal 60
Db 999 GCAATTGAAAACATATCAGCTTCGTGTCAACAGATTAAAGGACCCCTAAATAATCCTGGGCAT 1058
QY 61 CysGluThrThrProGlyValLysSerTyrSerGlyTyrValAspThrSerProGlu--- 79
Db 1059 -----GACCCAAATGTACACAGTACACGGGTACTTGGATGTGGAAAGACGAGGAC 1109
QY 80 SerHisThrPhePheThrPheGluAlaArgHisAsnProGluThrAlaProIleThr 99
Db 1110 AAGCATTTCTTTTGGACTTTTGAAGTAGAAGACGATCCTGCAAGGATCCGGTCATC 1169
QY 100 LeuTrpLeuAsnGlyGlyProGlySerAspSerIleuIleGlyLeuPheGluGluLeuGly 119
Db 1170 CTTTGGTTGAACGGGGTCCAGGTTGTTCTTCACTAAACCGGGCTGTTCTTTGAATTAGGA 1229
QY 120 ProCysHisValAsnSerThrPheAspAspTyrIleAsnProHisSerTipAsnGluVal 139
Db 1230 CCTCATCATCTGGACCTGATTTGAAACCCATCGGGAACCTTACTCTTGAACAGCAAT 1289
QY 140 SerAsnLeuLeuPheLeuSerGlnProLeuGlyValGlyPheSerTyrSerAspThrVal 159
Db 1290 GCCACCGGATCTTCTTTGACCGAGCTGTCAACGTTGGGTTCTCGTATTTC----- 1340
QY 160 AspGlySerIleAsnProValThrGlyValValGluAsnSerSerPheAlaGlyValGln 179
Db 1341 -----GGGTCTCTCA 1349
QY 180 GlyArgTyrProThrIleAspAlaThrLeuIleAspThrThrAsnLeuAlaAlaGluAla 199
Db 1350 GGT-----GTTTCCAACACTGTCCGCGCTGGTAAGAT 1382
QY 200 AlaTrpGluIleLeuGlnGlyPheLeuSerGlyLeuProSerLeuAspSerArgValGln 219
Db 1383 GTCATAAATCTTCTGGAGTTGTTCTTCGATCAGTTCCTCGAATACGTCAACAAGGGCCAA 1442
QY 220 SerLysAspPheSerLeuThrThrGluSerTyrGlyHisTyrGlyProAlaPhePhe 239
Db 1443 -----GATTTCCATCGCTGGGAATCCTACGCGCGCCATTATCATCCCTGTTTTCGCC 1496
QY 240 AsnHisPheTyrGluGlnAsnGluArgIleAlaAsnGlySerValAsnGlyValGlnLeu 259
Db 1497 TCTGAATTTTGTCTCAACAGGACAGA----- 1523
QY 260 AsnPheAsnSerLeuGlyIleIle-----AsnGlyIleIleAspGluAlaIleGlnAla 277
Db 1524 AACTTCAACTTAACCTCCGCTGTGATCGGAATGGCTCAGTCAACCCATTGACTCAGTAT 1583
QY 278 ProTyrTyrProGluPheAlaValAsnAsnThrTyrGlyIleLysAlaValAsnGluThr 297
Db 1584 AACTATTACGAACCAATGCGCTGTGTGAAGT-----GGGAACCATCTGTT----- 1631

```

```

QY 298 ValTyrAsnTyrMetLysPheAlaAsnGlnMetProAsnGlyCysGlnAspLeuIleSer 317
Db 1632 -----TTGCCCTCGAGGAATGCTCTGCTATGGAAGACTCTTTGGAA 1673
QY 318 ThrCysLysGlnThrAsnArgThrAlaLeuAlaAspTyrAlaLeu-----CysAlaGlu 335
Db 1674 CGTTGTTTGGGCTTGATCGAGTCGTCGTATGACTCGCAATCGGTCTCGTCTGTTGCCA 1733
QY 336 AlaThrAsnMetCysArgAspAsnValGluGlyProTyrTyrAlaPheAlaGlyArgGly 355
Db 1734 GCTACCAATTATTGTAATAACGCCCAATGGCTCTTTACCAACGT---ACGGCAGAAAC 1790
QY 356 ValTyrAspIleArgHisProTyrAspAsp-----ProThrProProSer 370
Db 1791 GTTTACCATATCAGGAAGGATTGTGAAGGTGGCAATTTGTGCTACCCCAACGTTACAAGAT 1850
QY 371 TyrTyrAsnLysPheLeuAlaLysAspSerValMetAspAlaIleGlyValAsnIleAsn 390
Db 1851 ---ATCGACGACTACTTTAAACACGAGACTACGTCAAAGAAGCTGTGGTGGGAGTTGAC 1907
QY 391 TyrThrGlnSer---AsnAsnAspValTyrTyrAlaPheGlnGlnThrGlyAspPheVal 409
Db 1908 CACTACGAATCCTGTAACTTCGATATCAACAGAAATTTCTGTGTTGGGATGATGGATG 1967
QY 410 TipProAsnPheIleGluAspLeuGluGluIleLeuAlaLeuProValArgValSerIleu 429
Db 1968 AAGCT---TACCACACCGCGGTAACAGATCTTTTGAATCAAGACCTTACCCATTCTGTA 2024
QY 430 IleTyrGlyAspAlaAspTyrIleCysAsnTrpPheGlyGlyGlnAlaValSerIleuAla 449
Db 2025 TATCGAGCGGATAAAGATTTTCATCTGTAACCTGTGGTAAATAAGGCGTGGACGGATGC 2084
QY 450 AlaAsnTyrSerGlnAlaAlaGlnPheArgSer-----AlaGlyTyrThrProLeu 466
Db 2085 TTGCCATGGAAGTAGTACGAGGAAGAAATTTGCAAGCCAAAGTACGTAAGTACGCTGCTTCT 2144
QY 467 LysValAsnGlyValGluTyrGlyGluThrArgGluTyrGlyAsnPheSerPheThrArg 486
Db 2145 ATCACCAGCAGAGGTCGCT---GGTGAAGTCAATCCTACAAGCAGCTTCAACCTATTGAGA 2201
QY 487 ValTyrGluAlaGlyHisGluValProTyrTyrGlnProIleAlaSerLeuGlnLeuPhe 506
Db 2202 GTCTTCAATGTGTGCCACATGGTTCCATTTCAGCTCCCTGAAACGCGCTTAAGTATGTT 2261
QY 507 AsnArgThrIlePheGly 512
Db 2262 AACGAATGGATCCACGGT 2279

```

Search completed: July 29, 2004, 13:38:07

Job time : 133 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 29, 2004, 11:31:29 ; Search time 702 seconds
(without alignments)
3876.414 Million cell updates/sec

Title: US-09-712-338-2
Perfect score: 2979
Sequence: 1 MRGYEFLVPLVAASWALP.....HTQSSVPLTATSSSSVGMA 555

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 3222919 seqs, 2451570024 residues

Total number of hits satisfying chosen parameters: 6445838

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09712338/runat_22072004_170322_10413/app_query.fasta_1.711
-DB=Published Applications_NA -QFMT=fastcap -SUFFIX=rmnp -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100
-NCPU=6 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USERS=US09712338@cgn 1.1 221 @runat_22072004_170322_10413
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-XGAPOP=6 -XGAPEXT=7 -XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No. Score Match Length DB ID Description

SUMMARIES

1	1882	63.2	1859	16	US-10-369-493-36581	Sequence 36581, A
2	812.5	27.3	1686	16	US-10-369-493-27042	Sequence 27042, A
3	636	21.3	1527	16	US-10-369-493-25156	Sequence 25156, A
4	588.5	19.8	2218	15	US-09-993-192-1	Sequence 1, Appli
5	588.5	19.8	2218	15	US-10-206-619-1	Sequence 1, Appli
6	539	18.1	1289	16	US-10-369-493-3620	Sequence 3620, A
7	530.5	17.8	1619	16	US-10-369-493-27598	Sequence 27598, A
8	527.5	17.7	1599	16	US-10-369-493-25613	Sequence 25613, A
9	479.5	16.1	1751	13	US-10-425-114-6406	Sequence 6406, Ap
10	475	15.9	1551	9	US-09-938-842A-683	Sequence 683, App
11	475	15.9	1551	11	US-09-938-842A-683	Sequence 683, App
12	473	15.9	1855	13	US-10-425-114-11176	Sequence 11176, A
13	468.5	15.7	2018	17	US-10-437-963-52572	Sequence 52572, A
14	464.5	15.6	1768	16	US-10-382-239A-1	Sequence 1, Appli
15	463	15.5	1913	13	US-10-424-599-59411	Sequence 59411, A
16	455	15.3	1755	17	US-10-437-963-53563	Sequence 53563, A
17	455	15.3	1858	13	US-10-424-599-30149	Sequence 30149, A
18	450	15.1	1732	13	US-10-424-599-71565	Sequence 71565, A
19	397.5	13.3	1696	17	US-10-437-963-57156	Sequence 57156, A
20	388.5	13.0	1897	17	US-10-437-963-49604	Sequence 49604, A
21	382	12.8	1761	13	US-10-425-114-27988	Sequence 27988, A
22	381.5	12.8	1326	13	US-10-425-114-9868	Sequence 9868, Ap
23	375.5	12.6	1503	17	US-10-437-963-97753	Sequence 97753, A
24	372	12.5	1886	13	US-10-424-599-26173	Sequence 26173, A
25	368.5	12.4	1514	13	US-10-425-114-29826	Sequence 29826, A
26	368	12.4	1890	13	US-10-425-114-26356	Sequence 26356, A
27	366.5	12.3	1949	10	US-09-971-429B-33	Sequence 33, Appli
28	361	12.1	1551	14	US-10-084-018-4	Sequence 4, Appli
29	360	12.1	2100	13	US-10-424-599-139251	Sequence 139251, A
30	359	12.1	1697	10	US-09-796-753-39	Sequence 586, App
31	359	12.1	1963	13	US-10-296-115-586	Sequence 163, App
32	359	12.1	2076	9	US-09-909-320-163	Sequence 163, App
33	359	12.1	2076	9	US-09-909-088B-163	Sequence 163, App
34	359	12.1	2076	9	US-09-905-291A-163	Sequence 163, App
35	359	12.1	2076	9	US-09-902-853-163	Sequence 163, App
36	359	12.1	2076	9	US-09-907-824-163	Sequence 163, App
37	359	12.1	2076	9	US-09-907-841-163	Sequence 163, App
38	359	12.1	2076	10	US-09-904-011-163	Sequence 163, App
39	359	12.1	2076	10	US-09-906-742-163	Sequence 163, App
40	359	12.1	2076	10	US-09-906-838-163	Sequence 163, App
41	359	12.1	2076	10	US-09-907-613-163	Sequence 163, App
42	359	12.1	2076	10	US-09-907-942-163	Sequence 163, App
43	359	12.1	2076	10	US-09-904-859-163	Sequence 163, App
44	359	12.1	2076	10	US-09-909-204-163	Sequence 163, App
45	359	12.1	2076	10	US-09-904-820-163	Sequence 163, App

ALIGNMENTS

RESULT 1

US-10-369-493-36581
; Sequence 36581, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 36581
; LENGTH: 1859
; TYPE: DNA
; ORGANISM: Aspergillus nidulans
US-10-369-493-36581

Alignment Scores:

Pred. No.: 9,52e-215 Length: 1859
 Score: 1882.00 Matches: 382
 Percent Similarity: 70.06% Conservative: 51
 Best Local Similarity: 61.81% Mismatches: 56
 Query Match: 63.18% Indels: 133
 DB: 16 Gaps: 8

US-09-712-338-2 (1-555) x US-10-369-493-36581 (1-1859)

QY 26 SerValGlyArgArgGlnLeuProLysAsnProThrGlyValLysThrLeuThrAla 45
 DB 19 TCITTTCAACAGGACAGCTACCGAGGACCAACTGGTGCAGAAATTCACAGACTCCC 78
 QY 46 AsnAsnValThrIleArgTyrLysGluProGlyAlaGluGlyValCysGluThrPro 65
 DB 79 AATGGTGTGACCATCCGATACAAGAGAGCTGGTAAAGAGGGTGTCTGTGAGACGACCCCT 138
 QY 66 GlyValLysSerTyrSerGlyTyrValAspThrSerProGluSerHisThrPhePheTrp 85
 DB 139 GGTGTCAACTCTTCTCGGATAGCTTATCTATCGCTGAGTCTCACACCTTCTTTTAC 198
 QY 86 PhePheGluAlaArgHisAsnProGluThrAlaProIleThrLeuTrpLeuAsnGlyGly 105
 DB 199 TTTCTTGAAGCTCGTCATGACCCGGCAACGCTCTATCACTCTGTGGTGAATGGTGGC 258
 QY 106 ProGlySerAspSerLeuIleGlyLeuPheGlu 116
 DB 259 CCCGGAAGTATTCGTTATCGGGCTTTTGA-AGGTTTGACTGCTGACTGGGTACA 317
 QY 117 -----GluLeuGlyProCysHisValAsnSerThr 126
 DB 318 AGAAGTTGATGACTGATGGGCTTCTTGTAGAACTGGGACCGTGGCTATCAGCGAGAAC 377
 QY 127 PheAspAspTyrIleAsnProHisSerTrpAsnGluValSerAsnLeuPheLeuSer 146
 DB 378 CTGACGACTAGTCAACCCATACCTCGTGGATGAAGTCTCAACCTTTGTTCATCTCG 437
 QY 147 GlnProLeuGlyVal 151
 DB 438 CAGCCATTGGGAGTCGGTATGAAGAAATTTGTCTTTAGTCTGTCGAGCTTCTCACTG 497
 QY 152 -----GlyPheSerTyrSerAspThrValAspGlySerIleAsnProValThrGlyVa 169
 DB 498 GSCCATAGAGATTTCTGATGCGGAAACAGAGAGGGTGGCTCAATCCCGTACCAGGTGT 557
 QY 169 lValGluAsnSerSerPheAlaGlyValGlnGlyArgTyrProThrIleAspAlaThrLe 189
 DB 558 GGTGAGGATAGTTCTTTGCTGGGTCCAGGTCGCTATCCCAAAATCAACGCTACCT 617
 QY 189 uile-----As 191
 DB 618 GAC-CGGTAATTTCTCAACAAACGTCGATGCTAGTTGTAGAGCTAAAGCTGTACTAGA 676
 QY 191 pThrThrAsnLeuAlaAlaGluAlaAlaTrpGluIleLeuGlnGlyPheLeuSerGlyLe 211
 DB 677 TACCAACCGAGCTCGCGGAGGCACTTTGGGAAGTCGTGCGAGGGCTTTCTCGGTGGT 736
 QY 211 uProSerLeuAspSerArgValGlnSerLysAspPheSerLeuTrpThrGluSerTyrGl 231
 DB 737 ACCTCAGTTGGACAGCACTATCAAGTCCAAAGACTTTAATTTATGAGCTGAGAGTTACGG 796
 QY 231 YGlyHisTyrGlyProAlaPhePheAsnHisPheTyrGluGlnAsnGluArgIleAlaAs 251
 DB 797 AGGACACTACGGTCCCGCATCTTCTCAATCACTTTTACGAGCAGAACCAAAAAATTCGCAA 856
 QY 251 nGlySerValAsnGlyValGlnLeuAsnPheAsnSerLeuGlyIleIleAsnGlyIleI 271
 DB 857 CGGAACACTGATGGTATGAGCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCA 916
 QY 271 eAspGluAlaIleGln----- 276

DB 917 CGATGAGGGGATTCAGGTACGGGATAGACAAGCGCGCTCAGCAATATCTCTCTGACTAGAA 976
 QY 277 -----AlaProTyrTyrProGluPheAlaValAsnAsnThrTyrGlyIleValAla 293
 DB 977 ACTAGGCACCTTTTACCTTACCTTGTGGGTGCACAAACACCTACGGCATCAAGC-CGTG 1035
 QY 294 -----ValAsnGluT 297
 DB 1036 AGTACTGTCTTTTTCAGCAAAATGGACGACATACTTACAAGCTGAAAAACAGGTTAAACGAGA 1095
 QY 297 hrValTyrAsnTyrMetLysPheAlaAsnGlnMetProAsnGlyCysGlnAspLeuIleS 317
 DB 1096 CAGTCTACAACTACATGAAATTTGCCAACAGATGGCAACCGATGCCAGATCAGATTG 1155
 QY 317 erThrCysLysGlnThrAsnArgThrAlaLeuAlaAspTyrAlaLeuCysAlaGluAlaT 337
 DB 1156 CCCTCTCAAGCAACAAACCGTCTCTTTTGGCGACTATGCTCTCTGCTCGAGGCGAA 1215
 QY 337 hrAsnMetCysArgAspAsnValGluGlyProTyrTyr----- 349
 DB 1216 CTAACATGTGCAGAGATAACGTT--GGTAGCTACTTCTACCTGAGACTCTGGGAAAATA 1272
 QY 350 -----AlaPheAlaGlyArgGlyV 356
 DB 1273 ATCAGCTGATTTATCTTCCACAGAGGACTCTCTACTACAGGCTACGA-GCAACTCGAGGCG 1331
 QY 356 alTyrAspIleArgHisProTyrAsp----- 364
 DB 1332 TATATGACTTTCGCCACCCGCAAAATGTAGCCGTTGATCCTAATCCAGGTTCTCTATAC 1391
 QY 365 -----AspProThrProProSerTyrTyrAsnLysPheIleAlaLysAspS 380
 DB 1392 TAACATAGAAAAGAGCCCAACACCCCGAACCTTCTCGTCCCTTCTCCAGCAAGACT 1451
 QY 380 erValMetAspAla-IleGlyValAsnIleAsnTyrThrGlnSerAsnAspValTyr 399
 DB 1452 GGGTCATGATGCTTCTTGGCGTGAACATCAACTACACCTCTTCCAAATAACGAAATTTAC 1511
 QY 400 TyrAlaPheGlnGlnThrGlyAspPheValTrpProAsnPheIleGluAspLeuGluGlu 419
 DB 1512 TATGCTTTCAGCAAAACCGGTGATTTCTTTTGGCCCAACTTCATTCAGGATCTCGAGAA 1571
 QY 420 IleLeuAlaLeuProValArgValSerLeuIleTyrGlyAspAlaAspTyrIleCysAsn 439
 DB 1572 ATCTCACCCTACCCAGTCCGCTGAGTCTGATCTATGTCGAGCAGCAGACTACATCTGCAAC 1631
 QY 440 TrpPheGlyGlyGlnAlaValSerLeuAlaAsnTyrSerGlnAlaAlaGlnPheArg 459
 DB 1632 TGGTTTCGAGGCCAAGCGCTCTCCCTCGCGCTCAAGTACCCGCAATCCCAAGCAGTTCAG 1691
 QY 460 SerAlaGlyTyrThrProLeuLysValAsnGlyValGluTyrGlyThrArgGluTyr 479
 DB 1692 GCCCGCGGATACACGCTTCTTGTGTGATGGGTGGAGTACGGAGAGACTCGAGATAT 1751
 QY 480 GlyAsnPheSerPheThrArgValTyrGluAlaGlyHisGluValProTyrTyrGlnPro 499
 DB 1752 GGCATTTCTCTCTCCGCTGCTATGAGCGCGCCACGAGGTTCTTACTATCAGCT 1811
 QY 500 IleAlaSerLeuGlnLeuPheAsnArgThrIlePheGlyTrpAspIle 515
 DB 1812 GCTGCGTCTGCTGAGCTTTTAAACCGCACTTTGTTGGGTGGGATATT 1859

RESULT 2

US-10-369-493-27042
 ; Sequence 27042, Application US/10369493
 ; Publication No. US2003023675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF

33	ProLysAsnProThrGlyValIleThrLeuThrThrAlaAsn-----AsnValThrIle	50
QY		
82	CCATATCAAGCTGTGGCTTGAAGACAGTCACCTTCCAACCGCAACATCAAGATC	141
DB		
51	ArgTyrLysGluProGlyAlaGluGlyValCysGluThrThr---ProGlyValLysSer	69
QY		
142	TCGTACAAGAACCC-----AAGGATGTCTGTACGACTCTTTCAGACGGCAGAACGAG	195
DB		
70	TyrSerGlyTyrValAspThrSerProGlu-----SerHisThrPhePheTrpPhePhe	87
QY		
196	TACACCGGCTGGTTCGTCCTCCGCGGAGTACCCAAACACACTCTTCTCTGTGTGTGT	255
DB		
88	GluAlaArgHisAsnProGluThrAlaProIleThrLeuTrpLeuAsnGlyGlyProGly	107
QY		
256	GGTGCCCGCGAG-----CCAACTCTCGCTGACATGATGCTCAACGGCGGGCCCGCG	309
DB		
108	SerAspSerLeuLeuGlyLeuPheGluGluLeuGlyProCysHisVal-----	123
QY		
310	TCGAGTTCCATGTTTGCTGTTCGCGGAGAATGGCCCTTGCAGGTCGTCGAGAGGGC	369
DB		
124	AsnSerThrPheAspAspTyrIleAsnProHisSerTrpAsnGluValSerAsnLeuLeu	143
QY		
370	GCGAGCCGCTGGAGCGCGCGCCAGAGAAATGGGGCTGGGATAGAGCCTCCAATATGTC	429
DB		
144	PheLeuSerGlnProLeuGlyValGlyPheSerTyrSerAspThrValAspGlySerIle	163
QY		
430	TTTGTGACACAGCCCAACCATGTGCGTTTCTCATACGATCTCCCAACCAACGGTCTCTG	489
DB		
164	AsnProValThrGlyValValGlu-----	171
QY		
490	GACTTGACGACTGGAGTGTGCGCTACCATGCATTCGCTGATAACCTGCTCCACG	549
DB		
172	-----AsnSerSerPheAla-----GlyValGlnGlyArgTyrProThrIleAsp	186
QY		
550	CTCTTCCTCAACGCGAGTTTCTCTCAAAACATCCAA-----	588
DB		
187	AlaThrLeuLeuAspThrThrAsnLeuAlaAlaGluAlaAlaTrpGluLeuGlnGly	206
QY		
589	---AACACGGCCAACAGCACTCAGAATGCTGCCATGGCGCTCTACCACTGCTCCAGGG	645
DB		
207	PheLeuSerGlyLeuProSer---LeuAspSerArgValGlnSerIlyAsnPheSerLeu	225
QY		
646	TTCTCTGACTTCTCCCGAGTAGCTCCCGGACAGCAGCCGCGTGGCGGTAACACCT	705
DB		
226	TrpThrGluSerTyrGlyHisTyrGlyProAlaPhePheAsnHisPheTyrGluGln	245
QY		
706	TTTACGGAAGATTATGCGGCCCACTACGGGCGCTGCTTTGCCGACACCTGGCAGAGGAG	765
DB		
246	AsnGluArgIleAlaAsnGlySerVal-----AsnGly	256
QY		

;; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES

;; FILE REFERENCE: 38-10(52052)B
 ;; CURRENT APPLICATION NUMBER: US/10/369,493
 ;; CURRENT FILING DATE: 2003-02-28
 ;; PRIOR APPLICATION NUMBER: US 60/360,039
 ;; PRIOR FILING DATE: 2002-02-21
 ;; NUMBER OF SEQ ID NOS: 47374
 ;; SEQ ID NO 25156
 ;; LENGTH: 1527
 ;; TYPE: DNA
 ;; ORGANISM: Saccharomyces cerevisiae
 US-10-369-493-25156

Alignment Scores:
 Pred. No.: 1.87e-65 Length: 1527
 Score: 636.00 Matches: 167
 Percent Similarity: 46.29% Conservative: 70
 Best Local Similarity: 32.62% Mismatches: 197
 Query Match: 21.35% Indels: 78
 DB: 16 Gaps: 14

US-09-712-338-2 (1-555) x US-10-369-493-25156 (1-1527)

QY 19 LeuProGlySerThrProAlaSerValGlyArgArgGlnLeu-----ProLys 34
 DB 127 CTTCACAGAACACAGAACCCCTTAATTTGGACCGTTGAATCAGATGATCCGCTG 186
 QY 35 AsnProThrGlyValLeuThrThrAlaAsnAsnValThrIleArgTyrLysGlu 54
 DB 187 TTTACAACTTTTATTTCTCTGTGACACAGATTACAGTTTGAGACCTTAGAACATAGAT 246
 QY 55 ProGlyAlaGluGlyValCysGluThrThrProGlyValLysSerTyrSerGlyTyrVal 74
 DB 247 CCTCTAACTAGGAATT-----GACACCGTAAACAAATGGTGGGTACATG 294
 QY 75 AspThrSerProGluSerHisThrPheThrPhePheGluAlaArgHisAsnProGlu 94
 DB 295 GACTATAGGATTCACAAACACTTTTCTACTGTGTTTGAAGTAGGACGATCCTGCT 354
 QY 95 ThrAlaProIleThrLeuTrpLeuAsnGlyGlyProGlySerAspSerLeuIleGlyLeu 114
 DB 355 AACGACCCAAATATTCTTTGGTTAAATGGTGACCTGTTGTTCTCTTTACTGGTGTG 414
 QY 115 PheGluGluLeuGlyProCysHisValAsnSerThrPheAspTyrIleAsnProHis 134
 DB 415 CTATTGGAATAGAGCCCTCATCAATTTGGCGCGATATGAAACCAATCCATCCCTAT 474
 QY 135 SerTrpAsnGluValSerAsnLeuLeuPheLeuSerGlnProLeuGlyValGlyPheSer 154
 DB 475 TCTTGGAAATAAACGCTTCAATGATCTCTTAGACACGCCACTCGGAGTCGGCTTTCC 534
 QY 155 TyrSerAspThrValAspGlySerIleAsnProValThrGlyValValGluAsnSerSer 174
 DB 535 TATGGTGTAT----- 543
 QY 175 PheAlaGlyValGlnGlyArgTyrProThrIleAspAlaThrLeuIleAspThrAsn 194
 DB 544 -----GAAAAGTCTCTCTACAAA 564
 QY 195 LeuAlaAlaGluAlaAlaTrpGluIleLeuGlnGlyPheLeuSerGlyLeuProSerLeu 214
 DB 565 TTAGCAGCAAGATCGGTACATTTTCTCTGGAATTTGTTTGAAGCTTTCTCTCATTTA 624
 QY 215 AspSerArgValGlnSerLysAspPheSerLeuTrpThrGluSerTyrGlyGlyHisTyr 234
 DB 625 -----CGCTCCACGATTTTCCATATGACGGGCAATCTTATGACGACATATAT 672
 QY 235 GlyProAlaPhePheAsnHisThrTyrGluGlnAsn---GluArgIleAlaAsnGlySer 253
 DB 673 ATCCCTCAAAATGCATGAGATCGTGTGCAAGAACCTTGAAAGA----- 717
 QY 254 ValAsnGlyValGlnLeuAsnPheAsnSerLeuGlyIleIleAsnGlyIleIleAspGlu 273

DB 718 -----ACGTTCAATTTAACTTCAGTTATGATTGTAATGGTATCAGACACCT 765
 QY 274 AlaIleGlnAlaProTyrTyrProGluPheAlaVal---AsnAsnThrTyrGlyIleLys 292
 DB 766 TTGATTCAAGCAGATTTATTATGAACCAATGCGATCGCGGAAAGGGGCTATCACCTGTT 825
 QY 293 AlaValAsnGluThrValTyrAsnTyrMetLysPheAlaAsnGlnMetProAsnGlyCys 312
 DB 826 CTCTCATCAGNAGATCGAGAAATAGTAAAGCTGCAGTCTG----- 870
 QY 313 GlnAspLeuIleSerThrCysLysGlnThrAsnArgThrAlaLeuAlaAspTyrAlaLeu 332
 DB 871 -----TGTCTAGGTTGAACAAGTTATGTTATGCTTCTTAATCAAGT 912
 QY 333 -----CysAlaGluAlaThrAsnMetCysArgAspAsnValGluGlyProTyrTyrAla 350
 DB 913 TTACATGATCATGTCGCACCTGCTTACTGTGACTCTGCACCTTTTGGAAACCGTACATTAAC 972
 QY 351 PheAlaGlyArgGlyValTyrAspIleArgHisProTyrAspAspPro----- 366
 DB 973 ---ACAGGACTCAACGCTCTATGACATTAGAGGGCCCTGTGAAGATATATAGTACTGATGGT 1029
 QY 367 -----ThrProSerTyrTyrAsnLysPheLeuAlaLysAspSerValMetAsp 383
 DB 1030 ATGTGTTATACAGGTCTCCGCTATGTCGACCATATATGAATTTTCTGAAAGTTCAAGAA 1089
 QY 384 AlaIleGlyValAsnIle---AsnTyrThrGlnSerAsnAsnAspValTyrTyrAlaPhe 402
 DB 1090 AGCTAGGCTCGACGTCATATTTCTGCTGTGTATGATGACGTCTTCAACCGATTT 1149
 QY 403 GlnGlnThrGlyAspPheValTrpProAsnPheIleGluAspLeuGluGluLeuAla 422
 DB 1150 TTGTTTACGGCGGATGGAAGTAAACCA---TTTCAACAATATATATGCTGAATTAATAA 1206
 QY 423 LeuProValArgValSerLeuIleTyrGlyAspAlaAspTyrIleCysAsnTrpPheGly 442
 DB 1207 CACAACATTCGGTATTATATATATGCGGTGATAAGGATTATTTGTAATTTGGCTGGGA 1266
 QY 443 GlyGlnAlaValSerLeuAlaAlaAsnTyrSerGlnAlaAlaGlnPheArgSerAlaGly 462
 DB 1267 AACCATGCTTGTCTCAATGAGTTGGAATGGATCAATAACGTAGTATCAGAGAAGATG 1326
 QY 463 TyrThrPro-----LeuLysValAsnGlyValGluTyrGlyGluThrArgGluTyrGly 480
 DB 1327 TTAAGACCATGGTCAAGTAAAGAACAGGTGAAGAGTTGGGACAAAGTCAAGAACTATGCG 1386
 QY 481 AsnPheSerPheThrArgValTyrGluAlaGlyHisGluValProTyrTyrGlnProIle 500
 DB 1387 CTTTTACCTTTTGGAGATATACGATCCGCTCATATGGTGGCTTATGATCAACCGGAG 1446
 QY 501 AlaSerLeuGlnLeuPheAsnArgThrIlePheGly 512
 DB 1447 GCAGTTTGAAGTAAAGTCAACAGTTGGATTTCGGT 1482

RESULT 4

US-09-993-192-1
 ; Sequence 1, Application US/09993192
 ; Publication No. US20020150983A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Korea Research Institute of Bioscience and Biotechnology
 ; APPLICANT: Dong Kook Pharmaceutical Co.
 ; APPLICANT: Rhee, Sangki
 ; APPLICANT: Choi, Hui-sung
 ; APPLICANT: Kang, Hyunah
 ; APPLICANT: Sohn, Junghoon
 ; APPLICANT: Bae, Junghoon
 ; APPLICANT: Kim, Moowoon
 ; APPLICANT: Agapronov, Michail
 ; TITLE OF INVENTION: Hansenula polymorpha mutants and process for the preparation of
 ; TITLE OF INVENTION: recombinant proteins using the same
 ; FILE REFERENCE: 4220-116 US
 ; CURRENT APPLICATION NUMBER: US/09/993,192
 ; CURRENT FILING DATE: 2001-11-14

APPLICANT: Bae, Junghoon

; APPLICANT: Kim, Moowong
 ; APPLICANT: Agaphonov, Michasel
 ; TITLE OF INVENTION: Hansenula polymorpha mutants and process for the preparation of
 ; TITLE OF INVENTION: Hansenula polymorpha mutants using the same
 ; FILE REFERENCE: 4220-116 US
 ; CURRENT APPLICATION NUMBER: US/10/206,619
 ; CURRENT FILING DATE: 2002-07-26
 ; PRIOR APPLICATION NUMBER: US/09/674,617
 ; PRIOR FILING DATE: 2001-01-03
 ; PRIOR APPLICATION NUMBER: PCT/KR00/00173
 ; PRIOR FILING DATE: 2000-03-04
 ; NUMBER OF SEQ ID NOS: 9
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 1
 ; LENGTH: 2218
 ; TYPE: DNA
 ; ORGANISM: Hansenula polymorpha DL1
 ; PUBLICATION INFORMATION:
 ; DATABASE ACCESSION NUMBER: U67174
 ; DATABASE ENTRY DATE: 1996-08-17
 ; RELEVANT RESIDUES: (1)..(2218)
 US-10-206-619-1

Alignment Scores:

Pred. No.:	1,69e-59	Length:	2218
Score:	588.50	Matches:	169
Percent Similarity:	45.61%	Conservative:	70
Best local Similarity:	32.25%	Mismatches:	200
Query Match:	19.75%	Indels:	86
DB:	15	Gaps:	20

US-09-712-338-2 (1-555) x US-10-206-619-1 (1-2218)

QY	19	LeuProGlySerThrProAlaSerValGlyArgArgGln-----Leu 32
DB	640	ATACCCTCAGAAACATGAAGCTCAAGTCAGAGCCGGCAGTCAAAATTACAAAGAACCTG 699
QY	33	ProLysAsnProThrGlyVallysThrLeuThrAlaAsnAsnValThrIleArgTyr 52
DB	700	CGGATTTTGGGACATTCATGTTCTCAATGAGAAAGTTC-TCCAACTACAGCTGAGGTT 758
QY	53	Lys-----GluProGlyAlaGluGlyValCysGluThrThrProGlyVallysSerTyr 70
DB	759	AGAAGACCGACCCGGGAGCATTTGGGACTGGACACACA-----AGACAGTAC 806
QY	71	SerGlyTyrValAspThrSerProGlu---SerHisThrPhePheThrPhePheGluAla 89
DB	807	TCTGGATCTTGATGTGGAGACGACGACGACGACGACGACGACGACGACGACGACGAC 866
QY	90	ArgHisAsnProGluThrAlaProIleThrLeuThrLeuAsnGlyGlyProGlySerAsp 109
DB	867	AGAAATGACCCGGTCAACGACCTGTGATCTCTGCTCTATCGCTCCAGTGGTCCAGATGCTCT 926
QY	110	SerLeuIleGlyLeuPheGluGluLeuGlyProCysHisValAsnSerThrPheAspAsp 129
DB	927	TCTTTGACTGGAATGCTTTTGGCTGGCTCTGCTCTATCGCTCCAGTCCAGATCTCAAGCCA 986
QY	130	TyrIleAsnProHisSerThrAsnGluValSerAsnLeuLeuPheLeuSerGlnProLeu 149
DB	987	ATCAACAACCCCATATTCGGGAAATCCCAATGCCACTGTGATTTCTCCACGACGCTGTC 1046
QY	150	GlyValGlyPheSerTyrSerAspThrValAspGlySerIleAsnProValThrGlyVal 169
DB	1047	AATGTTGGATTTCTCGTACTCT----- 1067
QY	170	ValGluAsnSerSerPheAlaGlyValGlnGlyArgTyrProThrIleAspAlaThrLeu 189
DB	1068	-----TCCAAGTCT 1076
QY	190	IleAspThrThrAsnLeuAlaAlaGluAlaAlaThrGluIleLeuGlnGlyPheLeuSer 209
DB	1077	GTTCCTAACACCGCTCGACGCTGTAAGAGCGTCTATGCTTCTTGGAGTTGTTTACACG 1136

QY	210	GlyLeuProSerLeuAspSerArgValGlnSerLysAspPheSerLeuTrpThrGluSer 229
DB	1137	CAATCCACACTTG-----CTGAGAACGACTTCCACATCCGCGGAGTCG 1184
QY	230	TyrGlyGlyHisTyrGlyProAlaPhePheAsnHisPheTyrGluGlnAsnGluArgIle 249
DB	1185	TACGGTGGTCATTACATCCAGTGTTCCTCCGAGATTCTCACCCTATGCTGACAGA--- 1241
QY	250	AlaAsnGlySerValAsnGlyValGlnLeuAsnPheAsnSerLeuGlyIleAlaAsnGly 269
DB	1242	-----TCTTCAACTCCTTCGGTGTGATTGGTAACGGT 1277
QY	270	IleIleAspGluAlaIleGlnAlaProTyrTyrProGluPheAlaValAsnAsnThrTyr 289
DB	1278	TTGACCGACCCACTTAACCATCTTACGAGAGATGGCATGCTCTACTGATGGT 1337
QY	290	GlyIleLysAlaValAsnGluThrValTyrAsnTyrMetLysPheAlaAsnGlnMetPro 309
DB	1338	GGCTATGAGCAA---CCCTGGACGAGTCTGAGTCGAAAGGAATGTTTGGAGACCTTGCT 1394
QY	310	AsnGlyCysGlnAspLeuIleSerThrCysLysGlnThrAsnArgThrAlaLeuAlaAsp 329
DB	1395	AGA---TGTTTGTCTATTGAATCATGCTACAGCTCGCAG-----TCT 1436
QY	330	TyrAlaLeuCysAlaGluAlaThrAsnMetCysArgAspAsnValGluGlyProTyrTyr 349
DB	1437	GTGTTCTCATGTGTCGGGCTCCCATCTACTGCAACACGACCACTTGGACCATCAAA 1496
QY	350	AlaPheAlaGlyArgGlyValTyrAspIleArgHisPro-----Tyr 363
DB	1497	AAG---ACCGGCAGAAACGCTCTACGACGTTAGAAGATGTGCGAGGAACTCTGTGCTAC 1553
QY	364	AspAspProThrProProSerTyrTyrAsnLysPheLeuAlaLysAspSerValMetAsp 383
DB	1554	AAAGAC-----ATGGAATACATGACCAATATTTGAACGAGACTTTGTCAAGGA 1604
QY	384	AlaIleGlyValAsnIleAsn---TyrThrGlnSerAsnAsnAspValTyrTyrAlaPhe 402
DB	1605	AAGGTTGGCTGGAGTTGACACTTACGAGTCGTGTAATTTTCGACGTCGAACAGAACTTC 1664
QY	403	GlnGlnThrGlyAspPheValTrpPro-----AsnPheIleGluAspLeuGlu 419
DB	1665	CTGTTTGTGCTGGATGGATGAAACCTTACCACAAAGACGTTATCAATCTGCTGGAGCAA 1724
QY	420	IleLeuAlaLeuProValArgValSerLeuIleTyr---GlyAspAlaAspTyrIleCys 438
DB	1725	-----GGTCTTCCCTGTC-----CTGATTACGAGGAGACAGGATTTTCATCTGC 1769
QY	439	AsnTrpPheGlyGlyGlnAlaValSerLeuAlaAlaAsnTyrSerGlnAlaAlaGlnPhe 458
DB	1770	AATTGGCTCGGAAACCAAGCCTGGTCCAATGAGCTCCCTTGGTCTGGACACGATGAATTC 1829
QY	459	ArgSerAlaGly---TyrThrProLeuLysValAsnGlyValGluTyrGlyGluThrArg 477
DB	1830	GAGTCCCGGACGCTCTACACCTCCTTGAAGGATGGCACTAAGCTGGCGAGGTCTCAAG 1889
QY	478	GluTyrGlyAsnPheSerPheThrArgValTyrGluAlaGlyHisGluValProTyrTyr 497
DB	1890	ANTGCTGGCAAGTTCACCTTCTAGTAATGTTGATGGAGGACACATGCTTCCATACGAC 1949
QY	498	GlnProIleAlaSerLeuGlnLeuPheAsnArgThrIlePheGlyTyrAspIleAlaGlu 517
DB	1950	CAGCGCTGAGGCTCTTGGCTATGCTCAATAGATGATAGCTGGT---GACTACTCCTTGT 2006
QY	518	GlyGlnLysLys 521
DB	2007	GGAAACCAAGAA 2018

RESULT 6

US-10-369-493-36920
 ; Sequence 36920, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:

; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; CURRENT FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; PRIOR FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 36920
 ; LENGTH: 1289
 ; TYPE: DNA
 ; ORGANISM: Aspergillus nidulans
 US-10-369-493-36920

Alignment Scores:
 Pred. No.: 6,14e-54 Length: 1289
 Score: 539.00 Matches: 149
 Percent Similarity: 44.86% Conservative: 69
 Best Local Similarity: 30.66% Mismatches: 171
 Query Match: 18.09% Indels: 98
 DB: 16 Gaps: 18
 US-09-712-338-2 (1-555) x US-10-369-493-36920 (1-1289)

QY 65 ProGlyValIysSerTyrSerGlyTyrValAspThrSerProGluSerHisThrPhePhe 84
 Db 10 CCCGATGCAAGCAATACACGGGTATCTGGATGACAAATGAAATGATAA-ACATTGGTT 68
 QY 85 -----Tip 85
 Db 69 TTACTGTAAGTGTATATTGCTGATGGGAGAGTAATATATAACGAGCTGAAGGG 128
 QY 86 PhePheGluAlaArgHisAsnProGluThrAlaProIleThrLeuTrpLeuAsnGlyGly 105
 Db 129 TTCTTTGAGTCAGCAAGACGCCCAAAACGATCCTGTGTTCTGTGTTGTAACGGTGGC 188
 QY 106 ProGlySerAspSerLeuIleGlyLeuPheGluGluLeuGlyProCysHisValAsnSer 125
 Db 189 CCTGGCTGTTCTCTCTCACTGAGACTCTTCATGAGCTTGGGCCAGCAGTAATGATGAG 248
 QY 126 ThrPheAspAspTyrIleAsnProHisSerTyrAsnGluValSerAsnLeuPheLeu 145
 Db 249 AACATCAAGCCGGTTTATACCCGATGCGTGGAACTCCCAACGCTCCCGTTATCTCTTG 308
 QY 146 SerGlnProLeuGlyValGlyPheSerTyrSerAspThrValAspGlySerIleAsnPro 165
 Db 309 GACCAGCCGTCACAGTTGCTACTCTTACAGC----- 341
 QY 166 ValThrGlyValValGluAsnSerSerPheAlaGlyValGlnGlyArgTyrProThrIle 185
 Db 341 ----- 341
 QY 186 AspAlaThrLeuIleAspThrThrAsnLeuAlaGluAlaTrpGluIleLeuGln 205
 Db 342 GGCAGTACTCTCAGCAGACTGTT---GCTGCTGGAAGAGACGTTTACGCCCTTGCTTACT 398
 QY 206 GlyPheLeuSerGlyLeuProSerLeuAspSerArgValGlnSerLysAspPheSerLeu 225
 Db 399 CTGTCTCTCAAGCAATTCCT-----GAGTATGCCGAGCAGGATTTCCACATC 446
 QY 226 TrpThrGluSerTyrGlyHisTyrGlyProAlaPheAsnHisPheTyrGluGln 245
 Db 447 GCTGGAGAAATCTATGTCGCCACTACATTCTCTGTTTACTTCTGAGTCTCTCTCAC 506
 QY 246 AsnGluArgIleAlaAsnGlySerValAsnGlyValGlnLeuAsnPheAsnSerLeuGly 265
 Db 507 CAGAAGCGC-----AACATCAACCTGAAGTCCGTCCTTA 539

QY 266 IleIleAsnGlyIleIleAspGluAlaIleGlnAlaProTyrTyrProGluPheAlaVal 285
 Db 540 ATCGCAACGGCTTGACCGGCTCTACCCAGTACGAATACTACCGCCCATGCGCTGT 599
 QY 286 AsnAsnThrTyrGlyIleLysAlaValAsnGluThrValTyrAsnTyrMetLysPheAla 305
 Db 600 GCGCAGGGT---GGCTATCTGCTGTGTGATGAG---AGCAGCTGC 641
 QY 306 AsnGlnMetProAsnGly-----CysGlnAspLeuIleSerThrCysLysGlnThr 322
 Db 642 CGCTCCATGGACAATGCTCTTGGTCGATGCCAGTCTATGATTGATCTTGTCTACACAGT 701
 QY 323 AsnArgThrAlaLeuAlaAspTyrAlaLeuCysAlaGluAlaThrAsnMetCysArgAsp 342
 Db 702 GAGAGTGCT-----TGGGTCTGTGTCGCCGCTCCATATATCTGTAAACAAC 746
 QY 343 AsnValGluGlyProTyrTyrAlaPheAlaGlyGlyValTyrAspIleArgHisPro 362
 Db 747 GCCTCTCTCGCTCTTATACAGCG---ACCGGACAGAACGTGTACGACGTTCGGGGCAG 803
 QY 363 TyrAspAspProThr-----ProProSerTyrTyrAsnLysPheLeuAla 377
 Db 804 TCGGAGCAGCAGAGTAACCTTTGTACAAAGGTATGGTTACGTTTCCGAATATCTGAC 863
 QY 378 LysAspSerValMetAspAlaIleGlyValAsnIleAsn---TyrThrGlnSerAsnAsn 396
 Db 864 AAGCTGAAGTCCGTGCGGCTGCGGTGCGGAGGTGTGATGTTTACGACTCTCGCAATTT 923
 QY 397 AspValTyrTyrAlaPheGlnThrGlyAspPheValTrpPro-----AsnPheIle 414
 Db 924 GACATCAACCGCACTCTCTCCATGGTGTGATGGAACCTTACACCGCTTGTGTC 983
 QY 415 GluAspLeuGluIleLeuAlaLeuProValArgValSerLeuIleTyr---GlyAsp 433
 Db 984 CTGGCATCTTTCAG-----CAGATACCTGTT-----CTCATCTATGCTGTGAT 1028
 QY 434 AlaAspTyrIleCysAsnTrpPheGlyGlyAlaValSerLeuAlaAlaAsnTyrSer 453
 Db 1029 GCGATTTTCATCTGCAACTGGCTCGGAAACCAAGGCTTGGAGCTTGAATGCGCT 1088
 QY 454 GlnAlaAlaGlnPheArgSerAlaGlyTyrThrProLeuLysVal----- 468
 Db 1089 GGCACACAGGAGTTCGCCGAGCTCCCATGAGAGATCTAAAGATCTGCGACAGAACAC 1148
 QY 469 AsnGlyValGluTyrGlyGluThrArgGluTyrGlyAsnPheSerPheThrArgValTyr 488
 Db 1149 ACTGCAAGAAAGATCGCCAGATCAAGCCCATGGCAACTTCACTTTCATGCGCTCTAT 1208
 QY 489 GluAlaGlyHisGluValProTyrTyrGlnProIleAlaSerLeuGlnLeuPheAsnArg 508
 Db 1209 GTGGCGGCCACATGTTCCCATGAGCCAGCTGAGGCACTCCCTTGATGTTCTTCAACCGC 1268
 QY 509 ThrIlePheGly---Tip 513
 Db 1269 TGCTTGGTGGCGAGTGG 1286

RESULT 7

US-10-369-493-27598
 ; Sequence 27598, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; CURRENT FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; PRIOR FILING DATE: 2002-02-21

; NUMBER OF SEQ ID NOS: 47374

; SEQ ID NO 27598

; LENGTH: 1619

; TYPE: DNA

; ORGANISM: Neurospora crassa

US-10-369-493-27598

Alignment Scores:

Pred. No.: 9.24e-53 Length: 1619
Score: 530.50 Matches: 154
Percent Similarity: 43.91% Conservative: 66
Best Local Similarity: 30.74% Mismatches: 187
Query Match: 17.81% Indels: 95
DB: 16 Gaps: 18

US-09-712-338-2 (1-555) x US-10-369-493-27598 (1-1619)

Qy 47 AsnValThrIleArgTyrLysGluProGlyAlaGluGlyValCysGluThrThrProGly 66
Db 295 AACCTGCCGCTCAAGTCGTCGATCTCCAGCTCGGCTCGACAAG-----342
Qy 67 ValLysSerTyrSerGlyTyrValAspThrSerPro---GluSerHisThrPhePheTrp 85
Db 343 GTCACGACGTACAGCGGTTACCTGATGACGAGGAGACGATAGCACCTGTTACTGTG- 401
Qy 85 -----85
Db 402 TAAGTCAATACGCTCGAATGATCTGTTAGTAAATACGGTGCAGTCGACGACGTGT 461
Qy 86 -----PhePheGluAlaArgHisAsnProGluThrAlaProIleThrLeuTrpLeuAsn 103
Db 462 CCAGGGTTCTTCGAGCTTCGCAACGATCCCAAGACGACCCGCTGTTTGTGGTTGAC 521
Qy 104 GlyGlyProGlySerAspSerLeuIleGlyLeuPheGluGluLeuGlyProCysHisVal 123
Db 522 GGTGCCCGCGGTCTCTTCGCTCACTGCTCTCTTCCTTGAGCTCGGCCCTTCTTCCATT 581
Qy 124 AsnSerThrPheAspTyrIleAsnProHisSerTrpAsnGluValSerAsnLeu 143
Db 582 GACAAGAGCTCAGGGTGTGACGACGATGATCGCTTGACACCAACGACGCGGTTATC 641
Qy 144 PheLeuSerGlnProLeuGlyValGlyPheSerTyrSerAspThrValAspGlySerIle 163
Db 642 TTCCTTGACCGCCGCTCAAGCTCGGTACTCGTACTCC-----680
Qy 164 AsnProValThrGlyValGluAsnSerSerPheAlaGlyValGlnGlyArgTyrPro 183
Db 680 -----680
Qy 184 ThrIleAspAlaThrLeuIleAspThrThrAsnLeuAlaGluAlaAlaTrpGluIle 203
Db 681 -----GGCAACGCGCTCAGCACTGTCGTCGCGCAAGATGTTATGCTCTC 731
Qy 204 LeuGlnGlyPheLeuSerGlyLeuProSerLeuAspSerArgValGlnSerLysAspPhe 223
Db 732 CTTACCCCTCTCTCCACCAGTTCC-----GAGTATGCCAAGCAGGACTTC 779
Qy 224 SerLeuTrpThrGluSerTyrGlyHisTyrGlyProAlaPhePheAsnHisPheTyr 243
Db 780 CACATTGTCGGCAGCTTATGCTGCTACTACCTCTCTTTCCTTCGAGATCTTG 839
Qy 244 GluGlnAsnGluArgIleAlaAsnGlySerValAsnGlyValGlnLeuAsnPheAsnSer 263
Db 840 TCCACAGGATCGC-----AACATCAATCTCAATCC 872
Qy 264 LeuGlyIleIleAsnGlyIleIleAspGluAlaIleGlnAlaProTyrTyrProGluPhe 283
Db 873 GTCCTTATGGTAATGGCTTGAATGATGATCTCTTACCCGATGACGACTACCGCCCGC 932
Qy 284 AlaValAsnAsnThrTyrGlyIleLysAlaValAsnGluThrValTyrAsnTyrMetLys 303
Db 933 GCCTCGCGCAAGGTGGCTACCCCGCGCTTCTCAGCGAGGAGGTGCGCGACATGGAC 992

Qy 304 PheAlaAsnGlnMetProAsnGlyCysGlnAspLeuIleSerThrCysLysGlnThrAsn 323
Db 993 -----AATGCCCTCCCGAGA--TGCAGTCTCTTATCCGTAAGTCTCTACGAGAGC--- 1040
Qy 324 ArgThrAlaLeuAlaAspTyrAlaLeuCysAlaGluAlaThrAsnMetCysArgAsn 343
Db 1041 -----GGCAGTGTCTGCTGCTGCTCTCTCGCGCATCTACTGTAAACAATCAG 1088
Qy 344 ValGluGlyProTyrTyrAlaPheAlaGlyArgGlyValTyrAspIleArgHisProTyr 363
Db 1089 TTCATCGCGCCCTACACGCGC---ACTGGTCAGAACCTCTATGACATCCCGGGAAGTGC 1145
Qy 364 AspAspProThrProProSerTyr-----TyrAsnLysPheLeuAlaLys 378
Db 1146 GAGGATGATAGCAACCTCTGCTATAGCGCTCTTGGCTGGATCAGGACTACTCTGAACCA 1205
Qy 379 AspSerValMetAspAlaIleGlyValAsnIle---AsnTyrThrGlnSerAsnAsp 397
Db 1206 AAGGATGTATGAGCGCTCTCGCGCTCGAGGTTGAGGCTATGAGAGCTGCAACTTTGAT 1265
Qy 398 ValTyrTyrAlaPheGlnThrGlyAspPheValTrpProAsnPheIleGluAspLeu 417
Db 1266 ATCAACCGCAACTTCTCTTCCAGGGTGAAGTGAAGAGCC---TTCATCGCCTCGTC 1322
Qy 418 GluGluIleLeu---AlaLeuProValArgValSerLeuIleTyr---GlyAspAlaAsp 435
Db 1323 CCCGCGCATCTCAAGGAGATCCCGTT-----CTCATCTATCGCGTGATCGCGAC 1373
Qy 436 TyrIleCysAsnTrpPheGlyGlyGlnAlaValSerLeuAlaAlaAsnTyrSerGlnAla 455
Db 1374 TTCATCTGCAACTGCTCGCAACAAGGCTTGGAGCGAGGCTCTTGTAGTGGCCCGCAAG 1433
Qy 456 AlaGlnPheArgSerAlaGlyTyrThrProLeuLysValAsnGlyVal-----GluTyr 473
Db 1434 AACGGCTTCAACAGAGCTGAGCTCGAGGACCTCTCTTCTCCCAAGGCTGACAGGATAC 1493
Qy 474 GlyGluThrArgGluTyrGlyAsnPheSerPheThrArgValTyrGluAlaGlyHisGlu 493
Db 1494 GGCAAGGTCAAGTCTCCGGAACCTTCACCTTCATGTCAGATCTACAGGCTGGCCACATG 1553
Qy 494 ValProTyrTyrGlnProIleAlaSerLeuGlnLeuPheAsnArgThrIlePheGly--- 512
Db 1554 GTGCCAATGGACCGCTGAGAACTCGTTCCTTCAACAGGTGTTGGCGGTGAA 1613
Qy 513 Trp 513
Db 1614 TGG 1616

RESULT 8

US-10-369-493-25613

; Sequence 25613, Application US/10369493

; Publication No. US20030233675A1

; GENERAL INFORMATION:

; APPLICANT: Cao, Yongwei

; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slater, Steven C.

; APPLICANT: Goldman, Barry S.

; APPLICANT: Chen, Xiaofeng

; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES

; FILE REFERENCE: 38-10(52052)B

; CURRENT APPLICATION NUMBER: US/10/369,493

; CURRENT FILING DATE: 2003-02-28

; PRIOR APPLICATION NUMBER: US 60/360,039

; PRIOR FILING DATE: 2002-02-21

; NUMBER OF SEQ ID NOS: 47374

; SEQ ID NO 25613

; LENGTH: 1599

; TYPE: DNA

; ORGANISM: Saccharomyces cerevisiae

US-10-369-493-25613

Alignment Scores:

```

Pred. No.: 2.07e-52 Length: 1599
Score: 527.50 Matches: 149
Percent Similarity: 42.97% Conservatives: 71
Best Local Similarity: 29.10% Mismatches: 195
Query Match: 17.71% Indels: 97
DB: 16 Gaps: 19

US-09-712-338-2 (1-555) x US-10-369-493-25613 (1-1599)

QY 33 ProLysAsnProThrGlyValIleThr-----LeuThrAlaAsnAsn 47
Db 244 CCAAAATTCCTGAAGCAATCAAAACGAAGAAAGACTGGGACTTTGTGGTCAAGAAATGAC 303
QY 48 ValThrIleArgTyr-----LysGluProGlyAlaGluGlyVal 60
Db 304 GCAATTGAATACTATCAGCTTCGTGTCAACAGATTAAAGACCTTAAATCCTGGGCAT 363
QY 61 CysGluThrThrProGlyValIleSerTyrSerGlyTyrValAspThrSerProGlu--- 79
Db 364 -----GACCCAAATGTCAACAGTACACGGGTTACTTGGATGTGGAAGACGAGGAC 414
QY 80 SerHisThrPheThrPheGluAlaArgHisAsnProGluThrAlaProIleThr 99
Db 415 AAGCATTTCTTTTGGACTTTTGAAGTAGAAACGATCCTCGAAAGGATCCGGTCAATC 474
QY 100 LeuTrpLeuAsnGlyProGlySerAspSerLeuIleGlyLeuPheGluGluGly 119
Db 475 CTITGGTTGAACGGGGTCCAGGTGTCTTCACTACCGGGCGTGTCTTTGAATTAGGA 534
QY 120 ProCysHisValAsnSerThrPheAspAspTyrIleAsnProHisSerTrpAsnGluVal 139
Db 535 CCTCATCCATTGGACCTGATTGTAACCCCATCGGAACCTTACTCTTGGAAACAGCAAT 594
QY 140 SerAsnLeuLeuPheLeuSerGlnProLeuGlyValGlyPheSerTyrSerAspThrVal 159
Db 595 GCCACCGTGATCTCTTGACACCGCTGCAACGTTGGGTTCTCGTATTC- 645
QY 160 AspGlySerIleAsnProValThrGlyValGluAsnSerSerPheAlaGlyValGln 179
Db 646 -----GGTCTCTCA 654
QY 180 GlyArgTyrProThrIleAspAlaThrLeuIleAspThrThrAsnLeuAlaAlaGluAla 199
Db 655 GGT-----GTTTCCAACACTGTCGCGCTGGTGAAGAT 687
QY 200 AlaTrpGluIleLeuGlnGlyPheLeuSerGlyLeuProSerLeuAspSerArgValGln 219
Db 688 GTCTATAACTTTGGAGTTGTTCTTCGATCAGTTCCTCGAATACGTCACAAAGGGCCAA 747
QY 220 SerLysAspPheSerLeuTrpThrGluSerTyrGlyHisTyrGlyProAlaPhePhe 239
Db 748 -----GATTTCACATCGCTGGGAATCTACCGCGCATACATCCCTGTTTGGCC 801
QY 240 AsnHisPheTyrGluGlnAsnGluArgIleAlaAsnGlySerValAsnGlyValGlnLeu 259
Db 802 TCTGAAATTTTGTCTCAAGGACAGA----- 828
QY 260 AsnPheAsnSerLeuGlyIleIle-----AsnGlyIleIleAspGluAlaIleGlnAla 277
Db 829 AACTTCAACTTAACCTCCGCTTGTATCGGAATGGGCTCACTGACCCATGACTCAGTAT 888
QY 278 ProTyrTyrProGluPheAla-----ValAsnAsnThrTyr 289
Db 889 AACTATTACCAACCAATGGCTGTGGTGAAGTGGCGAACCACTGTTTGGCCCTCGGAG 948
QY 290 GlyIleLysAlaValAsnGluThrValTyrAsnTyrMetLysPheAlaAsnGlnMetPro 309
Db 949 GAATGCTCTGCTATGGAAGACTCTTTG----- 975
QY 310 AsnGlyCysGlnAspLeuIleSerThrCysLysGlnThrAsnArgThrAlaLeuAlaAsp 329
Db 976 GAACGTTGTTGGCTTGTATCGATCGATCGTGTCTGATCGCA-----TCC 1020

```

```

QY 330 TyrAlaLeuCysAlaGluAlaThrAsnMetCysArgAspAsnValGluGlyProTyrTyr 349
Db 1021 GTCTGGTCTGTGTCCAGCTTACCAATTTATTTGTAATAACGCCCAATTCGCTCTTACCAG 1080
QY 350 AlaPheAlaGlyArgGlyValTyrAspIleArgHisProTyrAspAsp----- 365
Db 1081 CGT---ACCGGCAGAAAGTTTACGATATCAGGAAGGATGTGAAGTGGGCAATTTGTGC 1137
QY 366 ---ProThrProSerTyrTyrAsnLysPheLeuAlaLysAspSerValMetAspAla 384
Db 1138 TACCAACGGTTTACAAGAT---ATCCAGACTACTTAAACACGAGACTAGTCAAGAAGCT 1194
QY 385 IleGlyValAsnIleAsnTyrThrGlnSer---AsnAsnAspValTyrTyrAlaPheGln 403
Db 1195 GTCGGTGGGAGGTGACCACTACGAATCTCTGTAACCTTCGATATCAACAGAAATTTCTCG 1254
QY 404 GlnThrGlyAspPheValTrpProAsnPheIleGluAspLeuGluIleLeuAlaLeu 423
Db 1255 TTTGCGGTGATTGGATGAAGCT---TACCACACCGCCGTAACAGATCTTTTGAATCAA 1311
QY 424 ProValArgValSerLeuIleTyrGlyAspAlaAspTyrIleCysAsnTrpPheGlyGly 443
Db 1312 GACCTACCCATCTCTGGTATATGACAGCGATAAAGATTTCATCTGTAAGTGGTGGTAAAT 1371
QY 444 GlnAlaValSerLeuAlaAlaAsnTyrSerGlnAlaAlaGlnPheArgSer----- 460
Db 1372 AAGCGGTGGACGGATGCTTGGCCATGGAAGTACGACGAAGATTTGCAAGCAAAAGTA 1431
QY 461 AlaGlyTyrThrProLeuLysValAsnGlyValGluTyrGlyGluThrArgGluTyrGly 480
Db 1432 CGTAACCTGGACTCTTCTATCACCGCAGGTGCT---GGTGAAGTCAATTCCTACAAG 1488
QY 481 AsnPheSerPheThrArgValTyrGluAlaGlyHisGluValProTyrTyrGlnProIle 500
Db 1489 CACTTACCTATTGAGAGTCTTCAATGGTGGCCACATGTTCCATTTGACGTCCCTGAA 1548
QY 501 AlaSerLeuGlnLeuPheAsnArgThrIlePheGly 512
Db 1549 AAGCCTTAAGTATGTGTTAAAGAAATGATCCACGCT 1584

```

RESULT 9

```

US-10-425-114-6406
; Sequence 6406, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 6406
; LENGTH: 1751
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700563631_FLI
US-10-425-114-6406

```

Alignment Scores:

```

Pred. No.: 1.37e-46 Length: 1751
Score: 479.50 Matches: 141
Percent Similarity: 43.98% Conservatives: 93
Best Local Similarity: 26.50% Mismatches: 219
Query Match: 16.10% Indels: 79
DB: 13 Gaps: 18

```


QY 10 LeuProLeuValAlaAlaSerTyrAlaLeuProGlySerThrProAlaSerValGlyArg 29
 Db 196 TTGCTCTCACTGCTGGC-----GAGGACCGGGGATGTTGAACGG 237
 QY 30 ArgGlnLeuProLysAsn-----ProThrGlyValLysThrLeuThrAlaAsnAsn 47
 Db 238 AAGTTCGTATTCCTCCGAATATATCTGCCAGATGGTGGCCCTACTGTGCGACGATTTAGGTCA 297
 QY 48 ValThrIleArgTyrLysGluProGlyAlaGluGlyValCysGluThrThrProGlyVal 67
 Db 298 CATGCTGTTATTACAGCTCCGAAATCTCGTGGC----- 333
 QY 68 LysSerTyrSerGlyTyrValAspThrSerProGluSerHisThrPhePheThrPhePhe 87
 Db 334 -----GCAAGCATGTTCTACTTCTCTTC 357
 QY 88 GluAlaArgHisAsnProGluThrAlaProIleThrLeuTyrLeuAsnGlyGlyProGly 107
 Db 358 GAGTCACGG---AACAAGAGGATGCTCTCTGTGTGATTGTTGACGGAGGGCTGGA 414
 QY 108 SerAspSerLeuIleGlyLeuPheGluGluLeuGlyProCysHisValAsnSerThrPhe 127
 Db 415 TGTAGTAGTGAGTGGCTGTGTTCTATGAGAAATGTCCTTTCAAGATCACTAGTAACATG 474
 QY 128 AspAspTyrIleAsnProHisSerTyrAsnGluValSerAsnLeuLeuPheLeuSerGln 147
 Db 475 TCTCTGCTGGATGAGATGGATGGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 534
 QY 148 ProLeuGlyValGlyPheSerTyrSerAspThrValAspGlySerIleAsnProValThr 167
 Db 535 CCTGTGGAACTGTTTTCAGCTACACACAGATGAGTATGATGATGATGATGATGATGAT 594
 QY 168 GlyValValGluAsnSerSerPheAlaGlyValGlnGlyArgTyrProThrIleAspAla 187
 Db 595 GGAGTT----- 600
 QY 188 ThrLeuIleAspThrThrAsnLeuAlaAlaGluAlaAlaTyrGluIleLeuGlnGlyPhe 207
 Db 601 -----AGCAATGATCTTTATGATTTCTGCGAGCTTTC 633
 QY 208 LeuSerGlyLeuProSerLeuAspSerArgValGlnSerLysAspPheSerLeuTyrThr 227
 Db 634 TTTGCGGAGCACCCTAAGTTG-----GCAAAAAACGACTTTTATATTACTGGA 681
 QY 228 GluSerTyrGlyGlyHisTyrGlyProAlaPhePheAsnHisPheTyrGluIleAsnGlu 247
 Db 682 GAGTCATATGCTGGGCACTATATCCAGCTTTTCTCCGAGTCCATAAAGGAAACAAG 741
 QY 248 ArgIleAlaAsnGlySerValAsnGlyValGlnLeuAsnPheAsnSerLeuGlyIleIle 267
 Db 742 -----GCTAAT-----GAGGAGATTCATATTAACTTAAGGATTCGTTATTGGA 786
 QY 268 AsnGlyIleIleAspGluAlaIleGlnAlaProTyrTyrProGluPheAlaValAsnAsn 287
 Db 787 AATGGGCTTACAGATCTGCACTCCATACCAAGCTTATCTGACTACGCTTTGGAA--- 843
 QY 288 ThrTyrGlyIleLysAlaValAsnGluThrValTyrAsnTyrMetLysPheAlaAsnGln 307
 Db 844 ---ATGGTCTATACGCAAAAGAG-----CATGATCGCTTAGAAAAAGATT 888
 QY 308 MetProAsnGlyCysGlnAspLeuIleSerThrCysLysGlnThrAsnArgThrAla--- 326
 Db 889 GTCCCACTG---TGCAGACTCATCAATCAAGCTTTGTGGAATCTGATGCCACAACTTCTGT 945
 QY 327 LeuAlaAspTyrAlaLeuCysAlaGluAlaThrAsnMetCysArgAspAsnValGluGly 346
 Db 946 TTGGCATCATATCTGTCTGCAACTCTTTTTCAGTGGTGAATGATGATGATGATGATGAT 1005
 QY 347 ProTyrTyrAlaPheAlaGlyArgGlyValTyrAspIleArgHisPro----- 362
 Db 1006 GTAAACTAT-----TACGATCAGGAAGAGTGCCTGGGAAGT 1044

QY 363 -----TyrAspAspProThrProSerTyrTyrAsnLysPheLeuAlaLysAspSer 380
 Db 1045 TTGTGCTACGATTC-----TCAACACATGGAGAAATTTTGAATCTGCAATCT 1092
 QY 381 ValMetAspAlaIleGlyVal---AsnIleAsnTyrThrGlnSerAsnAsnAspValTyr 399
 Db 1093 GTGAGGAAGTCACTTGTTGGTGGGACATAGACTTTGTCTCTCGCAGTACTTCAGTCTAT 1152
 QY 400 TyrAlaPheGlnGlnThrGlyAspPheValTyrProAsnPheIleGluAsp----- 416
 Db 1153 CAGGCAATGCTGTAGAT-----TGGATGAGGAATCTCGAGTTGGAATCCC 1200
 QY 417 -----LeuGluGluIleLeuAlaLeuProValargValSerLeuIleTyr---GlyAsp 433
 Db 1201 ACGTCTTGGAGATGGAATCAGCCTC-----CTTGTATGTGAGAA 1245
 QY 434 AlaAspTyrIleCysAsnTyrPheGlyGlyGlnAlaValSerLeuAlaAlaAsnTyrSer 453
 Db 1246 TATGATCTTATCTGCACTGGCTCGTAACTCGAGGTGGTGAATGCAATGGAGTGTGCA 1305
 QY 454 GlnAlaAlaGlnPheArgSerAlaGlyTyrThrProLeuLysValAsnGlyValGluTyr 473
 Db 1306 GGGAAAAACAACTTTGGGGCGCTTAAAGAGTTCCGTTTATTGTGGATGGCAAGAGCA 1365
 QY 474 GlyGluThrArgGluTyrGlyAsnPheSerPheThrArgValTyrGluAlaGlyHisGlu 493
 Db 1366 GGCTTGTAAAGACTTATGAACTCACTCAGTTTCTCAAGGTGAGAGATGCAGGACACATG 1425
 QY 494 ValProTyrTyrGlnProIleAlaSerLeuGlnLeuPheAsnArg 508
 Db 1426 GTTCCATGAGCAGCCCTAAAGCTGCAITGAAATGCTGAAGCGA 1470

RESULT 11

US-09-938-842A-683
 ; Sequence 683, Application US/09938842A
 ; Publication No. US20040009476A9
 ; GENERAL INFORMATION:
 ; APPLICANT: Harper, Jeff
 ; APPLICANT: Kreps, Joel
 ; APPLICANT: Wang, Xun
 ; APPLICANT: Zhu, Tong
 ; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
 ; FILE OF INVENTION: SAME, AND METHODS OF USE
 ; FILE REFERENCE: SCRIP1300-3
 ; CURRENT APPLICATION NUMBER: US/09/938,842A
 ; CURRENT FILING DATE: 2001-08-24
 ; PRIOR APPLICATION NUMBER: US 60/227,866
 ; PRIOR FILING DATE: 2000-08-24
 ; PRIOR APPLICATION NUMBER: US 60/264,647
 ; PRIOR FILING DATE: 2001-01-16
 ; PRIOR APPLICATION NUMBER: US 60/300,111
 ; PRIOR FILING DATE: 2001-06-22
 ; NUMBER OF SEQ ID NOS: 5379
 ; SEQ ID NO 683
 ; LENGTH: 1551
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 US-09-938-842A-683

Alignment Scores:

Pred. No.: 3,91e-46 Length: 1551
 Score: 475.00 Matches: 152
 Percent Similarity: 42.72% Conservative: 68
 Best Local Similarity: 29.51% Mismatches: 189
 Query Match: 15.94% Indels: 106
 DB: 11 Gaps: 20

US-09-712-338-2 (1-555) x US-09-938-842A-683 (1-1551)

QY 10 LeuProLeuValAlaAlaSerTyrAlaLeuProGlySerThrProAlaSerValGlyArg 29
 Db 196 TTGCTCTCACTGCTGGC-----GAGGACCGGGGATGTTGTAACGG 237

Qy 30 ArgGlnLeuProIysAsn-----ProThrGlyValIysThrLeuThrAlaAsnAsn 47
 Db 238 AAGTTGCTATATCCCGAATATATCTTGGCGATGGTGGCCCTACTGTCGACGATTTAGGTCAT 297
 Qy 48 ValThrIleArgTyrIysGluProGlyValAlaGluGlyValCysGluThrThrProGlyVal 67
 Db 298 CATGCTGGTATTACAGCTCCCGAATCTCGTGGC----- 333
 Qy 68 LysSerTyrSerGlyTyrValAspThrSerProGluSerHisThrPhePheThrPhePhe 87
 Db 334 -----GCAAGCATGTTCTACTTCTTCTC 357
 Qy 88 GluAlaArgHisAsnProGluThrAlaProIleThrLeuThrLeuAsnGlyGlyProGly 107
 Db 358 GAGTCACCG--AACAAAGAGGATGCTCTCTGTTGATTTGTTGACGGGAGGCGCTGGA 414
 Qy 108 SerAspSerLeuIleGlyLeuPheGluGluLeuGlyProCysHisValAsnSerThrPhe 127
 Db 415 TGATAGTAGTGGTGGTGGTCTATGAGATGGTCTCTTCAAGATCACTAGTAACATG 474
 Qy 128 AspAspTyrIleAsnProHisSerTrpAsnGluValSerAsnLeuLeuPheLeuSerGln 147
 Db 475 TCTCTGTGCTGGATGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 534
 Qy 148 ProLeuGlyValGlyPheSerTyrSerAspThrValAspGlySerIleAsnProValThr 167
 Db 535 CCTGTGGAACCTGGTTCAGCTACACGACAGATAAAAGTATATCCGTCATGACGAAACT 594
 Qy 168 GlyValValGluAsnSerSerPheAlaGlyValGlnGlyArgTyrProThrIleAspAla 187
 Db 595 GGAGTT----- 600
 Qy 188 ThrLeuIleAspThrThrAsnLeuAlaGluAlaAlaTrpGluIleLeuGlnGlyPhe 207
 Db 601 -----AGCAATGATCTTTATGATTTCTGCGAGCTTTC 633
 Qy 208 LeuSerGlyLeuProSerLeuAspSerArgValGlnSerLysAspPheSerLeuThr 227
 Db 634 TTGCGGAGCACCTTAAGTTG-----GCAAAACGACCTTTATATTACTGGA 681
 Qy 228 GluSerTyrGlyGlyHisTyrGlyProAlaPhePheAsnHisPheTyrGluGlnAsnGlu 247
 Db 682 GAGTATATGCTGGGCACTATATCCAGCTTTGCTCCGAGTCCATAAAGAAACAAAG 741
 Qy 248 ArgIleAlaAsnGlySerValAsnGlyValGlnLeuAsnPheAsnSerLeuGlyIlele 267
 Db 742 -----GCTAAT-----GAGGAGTTCATATTAACTAAAGGGATTCGCTATTGA 786
 Qy 268 AsnGlyIleAlaAspGluAlaIleGlnAlaProTyrTyrProGluPheAlaValAsnAsn 287
 Db 787 AATGGGCTTACAGATCCTGACCTCAATACCAGCTATCCGACCTACCTGACCTTTGGAA--- 843
 Qy 288 ThrTyrGlyIleLysAlaValAsnGluThrValTyrAsnTyrMetIysPheAlaAsnGln 307
 Db 844 ---ATCGGCTTAATCAGCAAAAGAG-----CATGATCGCTTAGAAAAGATT 888
 Qy 308 MetProAsnGlyCysGlnAspLeuIleSerThrCysLysGlnThrAsnArgThrAla--- 326
 Db 889 GTCCCACTG---TGCGAACTATCAATCAAGCTTTTGTGGAACCTGATGGCAACAATCTCTGT 945
 Qy 327 LeuAlaAspTyrAlaLeuCysAlaGluAlaThrAsnMetCysArgAspAsnValGluGly 346
 Db 946 TTGGCATCATCTTGTCTGCAACTCTTGTTCAGTGGGTGTAATGATGATGATGATGATGATG 1005
 Qy 347 ProTyrTyrAlaPheAlaGlyArgGlyValTyrAspIleArgHisPro----- 362
 Db 1006 GTAAACTAT-----TACGACATCAGGAAGAAGTGCCTGGGAAGT 1044
 Qy 363 -----TyrAspAspProThrProProSerTyrTyrAsnIlyPheLeuAlaLysAspSer 380
 Db 1045 TTGTGCTACCATTC-----TCAACATGAGAGAAATTTTGTGAATCTGCAATCT 1092
 Qy 381 ValMetAspAlaIleGlyVal---AsnIleAsnTyrThrGlnSerAsnAspValTyr 399

Db 1093 GTGAGGAGTCACCTGGTGGTGGGACATAGACTTTGCTCTCTCGACTACTTCAGTCTAT 1152
 Qy 400 TyrAlaPheGlnGlnThrGlyAspPheValTrpProAsnPheIleGluAsp----- 416
 Db 1153 CAGCAATGCTTGTAGAT-----TGGATGAGGAATCTCGAGGTTGGAAATCC 1200
 Qy 417 -----LeuGluGluIleLeuAlaLeuProValArgValSerLeuIleTyr---GlyAsp 433
 Db 1201 ACGCTCTTGAAGATGAATCAGCTC-----CTTGTATGCTGGAGAA 1245
 Qy 434 AlaAspTyrIleCysAsnTrpPheGlyGlyGlnAlaValSerLeuAlaAsnTyrSer 453
 Db 1246 TATGATCTTATCTCAACTGCTGTAACCTCGAGTGGGTGAATGCAATGAGTGGTCA 1305
 Qy 454 GlnAlaAlaGlnPheArgSerAlaGlyTyrThrProLeuIysValAsnGlyValGluTyr 473
 Db 1306 GGGAAACAAACTTTGGGGCGCTAAAGAGTTCCGTTTATTTGGATGGCAAGAACGA 1365
 Qy 474 GlyGluThrArgGluTyrGlyAsnPheSerPheThrArgValTyrGluAlaGlyHisGlu 493
 Db 1366 GGCTTTGTAAGACTTATGAACAACCTCACTTCTCAAGTGCAGATGCGAGGACACATG 1425
 Qy 494 ValProTyrTyrGlnProIleAlaSerLeuGlnLeuPheAsnArg 508
 Db 1426 GTTCAATGGACCGCTAAAGCTGCAATGAATAATGCTGAAGCGA 1470

RESULT 12
 US-10-425-114-11176
 ; Sequence 11176, Application US/10425114
 ; Publication No. US20040034888A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jingdong
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kowalic, David K.
 ; APPLICANT: Screen, Steven E.
 ; APPLICANT: Tabaska, Jack E.
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
 ; FILE REFERENCE: 38-21(53313) B
 ; CURRENT APPLICATION NUMBER: US/10/425,114
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 73128
 ; SEQ ID NO 11176
 ; LENGTH: 1855
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: 701038279_FLI
 US-10-425-114-11176

Alignment Scores:
 Pred. No.: 9,03e-46 Length: 1855
 Score: 473.00 Matches: 133
 Percent Similarity: 45.06% Conservative: 77
 Best Local Similarity: 28.54% Mismatches: 182
 Query Match: 15.88% Indels: 74
 DB: 13 Gaps: 17

US-09-712-338-2 (1-555) x US-10-425-114-11176 (1-1855)

Qy 62 GluThrThrProGlyValIysSer-----TyrSerGlyTyr-----ValAspThr 76
 Db 359 GATTCGAGCCTTCGGTTGAGGATTTGGGTCCACCAGCTGGCTATTACCCCATTCAGCAT 418
 Qy 77 SerProGluSerHisThrPhePheThrPhePheGluAlaArgHisAsnProGluThrAla 96
 Db 419 TCACATGCCGCAAGGATGTTCTACTTTTCTCGAATCAAGCAATAGGAGGAGAT--- 475
 Qy 97 ProIleThrLeuThrLeuAsnGlyGlyProGlySerAspSerLeuIleGlyLeuPheGlu 116
 Db 476 CCTGTGCTAATTTGGTTGACTGGGGACCTGGGTGTAGCAGTGAATGGCTTTGTTTAT 535


```

QY 467 LysValAsnGlyValGluTyrGlyClnThrArgGluTyrGlyAsnPheSerPheThrArg 486
Db 1442 GTTGTGTGATGATCAGAAAGCGGATTAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAG 1501
QY 487 ValTyrGluAlaGlyHisGluValProTyrTyrGlnProIleAlaSerLeuGlnLeuPhe 506
Db 1502 GTCCATAATGCGGTCACATGTTCCATGACACGCCCAAGGCTGCATTAGAGATCGTG 1561
QY 507 AsnArgThrIlePheGly 512
Db 1562 AAGAAGTGGATTATGGA 1579

RESULT 13
US-10-437-963-52572
; Sequence 52572, Application US/10437963
; Publication No. US2004012343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 52572
; LENGTH: 2018
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_54856C.1
US-10-437-963-52572

Alignment Scores:
Pred. No.: 3,58e-45 Length: 2018
Score: 468.50 Matches: 151
Percent Similarity: 41.94% Conservative: 78
Best Local Similarity: 27.66% Mismatches: 205
Query Match: 15.73% Indels: 113
DB: 17 Gaps: 20

US-09-712-338-2 (1-555) x US-10-437-963-52572 (1-2018)
QY 8 SerValLeuProLeuValAlaAlaSerTrpAlaLeuProGlySerThrProAla----- 25
Db 168 TCGCGCTCCCGCCGCTCGACCTCCCGTCCCGCGGGGAGCTCCCGCCCGCTCGCGG 227
QY 26 -----SerValGlyArgArgGlnLeuProGlyAsnProThrGly 38
Db 228 CCGTCGCGCTCATCCGCGCGCTCAACCTCCACCGCGGAGCGCTCCCGCTCCCGCTCCA 287
QY 39 ValLysThrLeuThrThrAlaAsnValThrIleArgTyrLysGluProGlyAlaGlu 58
Db 288 GCGCGCGTGTATGGCGAGCTCCCG-----CCGGAACCTTAGTCGAGA 329
QY 59 Gly-----Val 60
Db 330 GGCCCATCCACCTCGCTCCATGGCCACCGGGAAGAGCGGCGCTGTCGGCTGAAGACC 389
QY 61 CysGluThrThrProGlyValLysSerTyrSerGlyTyrValAspThrSerProGluSer 80
Db 390 TCGGCCACACACCGGGTACTACCGCC-----TCCCAAC-ACC 427
QY 81 His-----ThrPhePheTrpPheGluAlaArgHisAsnPro---GluThrAla 96
Db 428 CACGACGCAAGGCTGTCTACTCTCTTCGAGTCGAGGGGAGCAAGGGGAGACGAC 487

```

97 ProIleThrLeuTrpLeuAsnGlyGlyProGlySerAspSerLeuIleGlyLeuPheGlu 116
 488 CCGCTGGTGTATCTGGCTGACGGGGCCCGGCTGCAGCAGGAGCTCGCGCTCTTCTAC 547
 117 GluLeuGlyProCysHisValAsnSerThrPheAspAspTyrIleAsnProHisSerTrp 136
 548 GAGAACGGGGCGCTTCACATCGCCGACACATGCTCTCGTGTGAGAACGACTTGGCTGG 607
 137 AsnGluValSerAsnLeuLeuPheLeuSerGlnProLeuGlyValGlyPheSerTrpSer 156
 608 GATCAGAGGTCAATCTGATATATGTGATCATGCCGCCACCGGTACTGGGTTCAGTCACAGC 667
 157 AspThrValAspGlySerIleAsnProValThrGlyValValcLuAsnSerSerPheAla 176
 668 TCCAATCCGGCGACACCCGCTCAGCATGAAGCTGGT---GTTAGCAACGACCTGTATGCC 724
 177 GlyValGlnGlyArgTyrProThrIleAspAlaThrLeuIleAspThrThrAsnLeuAla 196
 724 ----- 724
 197 AlaGluAlaAlaTrpGluIleLeuGlnGlyPheLeuSerGlyLeuProSerLeuAspSer 216
 725 -----TTCTTCAGGCTCTTCTCACTGAGCACCCGAAATTT----- 760
 217 ArgValGlnSerIlyAspPheSerLeuTrpThrGluSerTyrGlyGlyHisTyrGlyPro 236
 761 -----GCTAAAAATGATTTCTATATCTATCTGCGGAAATCGTATCTCTGCACACTATATTCCT 814
 237 AlaPhePheAsnHisPheTyrGluGlnAsnGluArgIleAlaAsnGlySerValAsnGly 256
 815 GCCTTTCCAAAGTCGAGTGTCACAGGGGAAACAAG-----AACGTGAGGCG 859
 257 ValGlnLeuAsnPhe-AsnSerLeuGlyIleIleAsnGlyIleIleAspGluAlaIleG 276
 860 ATTCAATCAATTTGAGGGGTTTGTCTATTGGCAATGGTTAACGGATCCACGATACA 919
 276 nAlaProTyrTyrProGluPheAlaValAsnAsnThrTyrGlyIleIleAlaValAsnG 296
 920 ATCAAGGCATACACTGATTACTCATTTGGAT-----ATGGGTTAATTACAAA 967
 296 uThrValTyrAsnTyrMetIlyPheAlaAsnGlnMetProAsnGlyCysGlnAspLeu 316
 968 ATCAAGTTTAACAGGATT-----AACAAATTTGTTCCAACTGTGAACTTGCTAT 1018
 316 eSerThrCysIlysglnThrAsnArg---ThrAlaLeuAlaAspTyrAlaLeuCysAlaG 335
 1019 CAAGCTCTGGTACCTCTGGCATATATCTTGCCTGGCGCATGTTGTCGTCG----- 1073
 335 uAlaThrAsnMetCysArgAspAsnValGluGlyProTyrTyrAlaPheAlaGlyArg 355
 1074 -----AATCTAATATCTCTTCCATCGAG-----ACAATAATTGGGAAGAAAA 1117
 355 yValTyrAspIleArgHisPro-----TyrAspAspProThrProPr 369
 1118 TTATTATGACATCAGGAAGCATGTGTGGGAGCCTATGCTATGCAC-----CT 1165
 369 oSerTyrTyrAsnIlyPheLeuAlaLysAspSerValMetAspAlaIleGlyVal---As 388
 1166 CTCTAATATGGAGAAGTTTCTCCAAATGAATCAATCAGTCAGAGAGCCCTAGGTGTGAGA 1225
 388 nIleAsnTyrThrGlnSerAsnAsnAspValTyrTyrAlaPheGlnGlnThrGlyAspPh 408
 1226 CATACAGTTTGTATCATGCAGCCCAACTGTATATCAGGCCATGCTG-----CT 1273
 408 eValTrpProAsnPheIleGlu---AspLeuGluGluIleuAlaLeuProValArgVa 427
 1274 AGATTGGATGAGAACCTTGAAGTTGGAATCCCTGAACTCCTTGGAAATGCATCAAAAGT 1333
 427 lSerLeuIleTyrGlyAspAlaAspTyrIleCysAsnTrpPheGlyGlyGlnAlaValSe 447
 1334 GTTGATTTATGCTGGAGAGTATGATCTCATATGCACTGTGTAGGAACCACTCAAGATGGGT 1393
 447 rLeuAlaAlaAsnTyrSerGlnAlaAlaGlnPheArgSerAlaGlyTyrThrProLeu 467

```

Db      1394 GAATTCATGGAAGTCTGGAAGGAGCGCTTTGTGTCCTCATCTGAGGAACCTTCAC 1453
Qy      467 sValAsnGlyValGLuTyrGlyGluThrArgGluTyrGlyAsnPheSerPheThrArgVa 487
Db      1454 AGTTGATGAAAAGAACCGGTATTCTGAAAAGACTATGGACCTTTGAGTTCTTGAAGGT 1513
Qy      487 lTyrGluAlaGlyHisGluValProTyrTyrGlnProIleAlaSerLeuGlnLeuPheAs 507
Db      1514 CCACGATGTGGCCACATGGTACCAATGGATCAACGGAAGGTGGCTCTGGAGATGCTGAT 1573
Qy      507 nArgThrIlePheGly 512
Db      1574 GAGGTGACCTCAGGA 1589

RESULT 14
US-10-382-239A-1
; Sequence 1, Application US/10382239A
; Publication No. US20030219511A1
; GENERAL INFORMATION:
; APPLICANT: Societe des Produits Nestle
; TITLE OF INVENTION: No. US20030219511A1el Carboxypeptidase of Cocoa
; FILE REFERENCE: 88265-7006
; CURRENT APPLICATION NUMBER: US/10/382,239A
; CURRENT FILING DATE: 2003-03-05
; PRIOR APPLICATION NUMBER: EP 01116407.6
; PRIOR FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: PCT/EP02/07162
; PRIOR FILING DATE: 2002-06-28
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 1
; LENGTH: 1768
; TYPE: DNA
; ORGANISM: Theobroma cacao
US-10-382-239A-1

Alignment Scores:
Pred. No.:      8,74e-45      Length: 1768
Score:          464.50      Matches: 153
Percent Similarity: 43.59%      Conservative: 92
Best Local Similarity: 27.22%      Mismatches: 210
Query Match:    15.59%      Indels: 107
DB:             16      Gaps: 23

US-09-712-338-2 (1-555) x US-10-382-239A-1 (1-1768)
Qy      4 TyrGluPhe---LeuSerValLeuProLeuValAlaAlaSerTrpAla----- 18
Db      46 TACCGGTTTCTGTGTTCCCTTCCTTCCTCATTTCCATCTCCTCCGCGCGCGCTTCCTCC 105
Qy      19 -----LeuProGlySerThrProAlaSerValGlyArgArgGlnLeu 32
Db      106 TTCTTACAGCGCGCGCACTCGGAGGATCAAGTTTCCCTCGATACATACGGAAGAGTTG 165
Qy      33 -----ProLysAsnProThrGlyValLysThrLeuThrThrAla 45
Db      166 ATAAGGAGGTGAATTTGTTTCTTCTAAGGAGGAAGTCAACGTC-----GTTGATGGA 216
Qy      46 AsnAsnValThrIle-----ArgTyrLysGluProGly 56
Db      217 GGCACAGGTTTCTTACCGGAGGATTCGAGGTTGGTGGAAGACGGTTCAAGTTCCCGAAT 276
Qy      57 AlaGluGlyValCysGluThrThrProGlyValLysSer-----Tyr 70
Db      277 TTG-----GCGGTGCTGTGGTGGGTTTCCGTGTGAGGATTTGGGTCATCAT 321
Qy      71 SerGlyTyr-----ValAspThrSerProGluSerHisThrPhePheThrPhePheGlu 88
Db      322 GCTGGTTATTACAAGCTAGCTAATTCATCATGATGCCAGAATGTTCTATTCTCTTTGAA 381
Qy      89 AlaArgHisAsnProGluThrAlaProIleThrLeuTrpLeuAsnGlyGlyProGlySer 108

```

Db 382 TCACGA---AATAGCAAAAGGACCCCTGTTGTAATCTGTTGACTGGAGGCCAGGGTGT 438
Qy 109 AspSerLeuIleGlyLeuPheGluLeuGlyProCysHisValAsnSerThrPheAsp 128
Db 439 AGTAGTGAATGGCTTTGTTTATGAAATGGTCTTTTACCATTTGCTGGAACATGTCT 498
Qy 129 AspTyrIleAsnProHisSerTyrAsnGluValSerAsnLeuPheLeuSerGlnPro 148
Db 499 CTTATTGGATCAGATGGTGGGACATGCATCAACCTTCTGTATGTGGACCAACCC 558
Qy 149 LeuGlyValGlyPheSerTyrSerAspThrValAspGlySerIleAsnProValThrGly 168
Db 559 ATTGGTACCGCTTTAGTTATAGT----- 582
Qy 169 ValValGluAsnSerSerPheAlaGlyValGlnGlyArgTyrProThrIleAspAlaThr 188
Db 583 -----TCTGATAGAGGACATTCGTCAATATGAAGATGAAGTTAGCAACGAC 630
Qy 189 LeuIleAspThrThrAsnLeuAlaGluAlaAlaTyrGluIleLeuGlnGlyPheLeu 208
Db 631 CTATATGAC-----TTCCTACAGCATTCCTTT 657
Qy 209 SerGlyLeuProSerLeuAspSerArgValGlnSerTyrAspPheSerLeuThrGlu 228
Db 658 GCTGAACACCCCTGAGTTT-----GAAAAGAATGACTTTTATATAACTGGAGAA 705
Qy 229 SerTyrGlyGlyHisTyrGlyProAlaPhePheAsnHisPheTyrGluGlnAsnGluArg 248
Db 706 TCATATGCTGGCCTACATTCCTGCTCCCGAGTCCACCAAGGAACAA--- 762
Qy 249 IleAlaAsnGlySerValAsnGlyValGlnLeuAsnPheAsnSerLeuGlyIleIleAsn 268
Db 763 -----GCTAAGATGGAAATTCATATAACCTTAAGGATTTGCTATTGCTAAT 810
Qy 269 GlyIleIleAspGluAlaIleGlnAlaProTyrTyrProGluPheAlaValAsnAsnThr 288
Db 811 GGCCTGACTGACCCCTGCAATCCAGTATAAAGCTTACACAGATATTGCTTTGGAC----- 864
Qy 289 TyrGlyIleLysAlaValAsnGluThrValTyrAsnTyrMetLysPheAlaAsnGlnMet 308
Db 865 -----ATGGGGGTAATTAAGAACTGCTGACTACAACTCGTATC-----AACAAGCTG 909
Qy 309 ProAsnGlyCysGlnAspLeuIleSerThrCysLysGlnThrAsnArg---ThrAlaLeu 327
Db 910 GTTCCAGTTGTTGAAATGGCAATAAAGCTTTGTGCACTGATGGCACAACTCTTGTGATG 969
Qy 328 AlaAspTyrAlaLeuCysAlaGluAlaThrAsnMetCysArgAspAsnValGluGlyPro 347
Db 970 GCTTCATATTTTGTCTGC-----AATGCCATATTCACTGGC 1005
Qy 348 TyrTyrAlaPheAlaGly---ArgGlyValTyrAspIleArgHisProTyrAspAspPro 366
Db 1006 ATCATGGCACTGTGTGGCGATACAAATTAATACGACATTAAGCAAAATGTGAAGGGAGC 1065
Qy 367 ThrProProSerTyrTyrAsn-----LysPheLeuAlaLysAspSerValMetAspAla 384
Db 1066 CTTTGCTATGACTCTCAAAATGAGGACATTTCTGAACAGGAAATCTGTAGGATGCC 1125
Qy 385 IleGlyVal---AsnIleAsnTyrThrGlnSerAsnAsnAspValTyrTyrAlaPheGln 403
Db 1126 CTTGGAGTTGGGATGATTGACTTTGTGCTGTCGATCTACAGTGTATCAGGCCCATGCTG 1185
Qy 404 GlnThrGlyAspPheValTyrProAsnPheIleGlu---AspLeuGluGluIleLeuAla 422
Db 1186 GTTGAC-----TGGATGAGGAATCTTGAAGTTGGCAATTCCTGCTCTCTCTGAG 1233
Qy 423 LeuProValArgValSerLeuIleTyrGlyAspAlaAspTyrIleCysAsnTyrPheGly 442
Db 1234 GATGGTCTCAGCTCTTGATATGCTGGGAATATGATCTCATCTGCAACTGCTGCTGGC 1293
Qy 443 GlyGlnAlaValSerLeuAlaAsnTyrSerGlnAlaAlaGlnPheArgSerAlaGly 462
Db 1294 AATTGAGATGGTCTCATGCAATGGAATGGTCTGTCAGAGGAGTATTGTAGCATCTCCT 1353

Qy 463 TyrThrProLeuLysValAsnGlyValGluTyrGlyGluThrArgGluTyrGlyAsnPhe 482
Db 1354 GAGGTTCCTTTTTCGTTTGATGGCTCAGAAGCAGGAGTCTTTGAGAATCTCATGGACCTCTT 1413
Qy 483 SerPheThrArgValTyrGluAlaGlyHisGluValProTyrTyrGlnProIleAlaSer 502
Db 1414 GGTTCCTAAAGTTTCAGATGACGATGACATGTTTCTATGGACCAAGGACGCA 1473
Qy 503 LeuGlnLeuPheAsnArgThrIlePheGlyTyr-----AspIleAlaGluGlyGlnLys 520
Db 1474 TTGGAGATGCTGAAGCGGTGGACTAAGGTACATTATCTGAAGCTGCGGATTACAGAGAA 1533
Qy 521 LysIle 522
Db 1534 TTGGTT 1539
RESULT 15
US-10-424-599-59411
; Sequence 59411, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 59411
; LENGTH: 1913
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_2465C.1
US-10-424-599-59411
Alignment Scores:
Pred. No.: 1,5e-44 Length: 1913
Score: 463.00 Matches: 124
Percent Similarity: 44.93% Conservative: 80
Best Local Similarity: 27.31% Mismatches: 180
Query Match: 15.54% Indels: 70
DB: 13 Gaps: 15
US-09-712-338-2 (1-555) x US-10-424-599-59411 (1-1913)
Qy 70 TyrSerGlyTyr-----ValAspThrSerProGluSerHisThrPhePheThrPhePhe 87
Db 434 CATGACGGCTATTATCGCTTCCTCGTTCCAAAGCTGCAAGGATGTTCTACTTCTCTTT 493
Qy 88 GluAlaArgHisAsnProGluThrAlaProIleThrLeuThrLeuAsnGlyGlyProGly 107
Db 494 GAATCTCGA---AGCATAAGAATGATCCTGTCTATATGGCTGATGGAGGACCAAGCA 550
Qy 108 SerAspSerLeuIleGlyLeuPheGluGluLeuGlyProCysHisValAsnSerThrPhe 127
Db 551 TGCAGCAGTGAATGATGCTTTTACGAAATGGTCTCTTTTCAGCTTACCAAGAACTTG 610
Qy 128 AspAspTyrIleAsnProHisSerTyrAsnGluValSerAsnLeuLeuPheLeuSerGln 147
Db 611 TCTCTTGTTGGAATGACTATGCTGGACAGCATCAAAATATATATATTATTGTAGACCAA 670
Qy 148 ProLeuGlyValGlyPheSerTyrSerAspThrValAspGlySerIleAsnProValThr 167
Db 671 CCCACTGGACGGGTTTCAGCTATCATCTCTGATGAGAGTGCACATTCGCCATGATGAAGAG 730
Qy 168 GlyValValGluAsnSerSerPheAlaGlyValGlnGlyArgTyrProThrIleAspAla 187
Db 731 GCGTT----- 736

```
QY 188 ThrLeuIleAspThrThrAsnLeuAlaAlaGluAlaAlaIleLeuGlnGlyPhe 207
Db      ::::: |||||
737 -----AGCAATGATTTGTATGACTTCTTGAGGCAATTT 769
QY 208 LeuSerGlyLeuProSerLeuAspSerArgValGlnSerLysAspPheSerLeuTrpThr 227
Db      ::::: |||||
770 TTCAAGGAGCACCT-----CAATTCACTAAGATGACTTTTATATTACTGA 817
QY 228 GluSerTyrGlyGlyHisTyrGlyProAlaPheAsnHisPheTyrGluGlnAsnGlu 247
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
818 GAGTCATATGCTGGACATTCACATTCAGCTCTTGATCCAGGGTTCACCAAGAAACAAA 877
QY 248 ArgIleAlaAsnGlySerValAsnGlyValGlnLeuAsnPheAsnSerLeuGlyIleIle 267
Db      ::::: |||||
878 -----GCAAAAGAGGAATTCATATAAATCTAAAGGATTTGCTATTGGT 922
QY 268 AsnGlyIleIleAspGluAlaIleGlnAlaProTyrTyrProGluPheAlaValAsnAsn 287
Db      ||||| ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
923 AATGGTTAAACCAATCCTGAAATTCAGTACACAGGCGGTATACAGACTACGCATTAGACAGG 982
QY 288 ThrTyrGlyIleLeuAlaValAsnGluThrValTyrAsnTyrMetLysPheAlaAsnGln 307
Db      ||||| ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
983 -----GGACTA-----ATTAAAGAGCGTGAATATACAGTATC-----AACAG 1021
QY 308 MetProAsnGlyCysGlnAspLeuIleSerThrCys---LysGlnThrAsnArgThrAla 326
Db      ::::: |||||
1022 TTGATCCCAACCATGCAAGCAGGCCATAGAACTTCGGAACTGAAGGTGGAGAACATGC 1081
QY 327 LeuAlaAspTyrAlaLeuCysAlaGluAlaThrAsnMetCysArgAspAsnValGluGly 346
Db      ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1082 GTATCTTCATTATATGTTTGCATAAGATATCAATCGGATCATGACCATTCGTCTGATGAT 1141
QY 347 ProTyrTyrAlaPheAlaGlyArgGlyValTyrAspIleArgHisPro----- 362
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1142 GTTATTAC-----TATGACATTAGAAAGAAATGTGTGGAGTT 1180
QY 363 -----TyrAspAspProThrProSerTyrTyrAsnLysPheLeuAlaLysAspSer 380
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1181 CRTGTGTATGACTTC-----TCTGTGATGGAGGATTTCTTAAACGAGAAGACA 1228
QY 381 ValMetAspAlaIleGlyVal---AsnIleAsnTyrThrGlnSerAsnAsnAspValTyr 399
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1229 GTGAGAGATGCTTTAGGTGTTGGGACTTGGACTTCGTTTCATGCAGTAGACAGCTGAT 1288
QY 400 TyrAlaPheGlnGlnThrGlyAspPheValTyrProAsnPheIleGlu---AspLeuGlu 418
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1289 AGTGTATGATCAAGAC-----TGGATGAGAAATCTGGAAGTTGCTATTCCTCC 1336
QY 419 GluIleLeuAlaLeuProValArgValSerLeuIleTyrGlyAspAlaAspTyrIleCys 438
Db      ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1337 ACACCTCTTGAGGAAGGAATCAAGTGTGTGTATGCTGGAGAGAGAGATCTCATATGC 1396
QY 439 AsnTrpPheGlyGlyGlnAlaValSerLeuAlaAlaAsnTyrSerGlnAlaIleGlnPhe 458
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1397 AATTGCTTGGGAATCAAGGTGGTTCAAGCAATGGAGTGGTCTCGTCAAAACAAATTT 1456
QY 459 ArgSerAlaGlyTyrThrProLeuLysValAsnGlyValGluTyrGlyLeuThrArgGlu 478
Db      ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1457 GGGGCGTGTGTACAGTCCCATTTTGGTTGATGTTGCTGAGCAGGACGCTGAAAGC 1516
QY 479 TyrGlyAsnPheSerPheThrArgValTyrGluAlaGlyHisGluValProTyrTyrGln 498
Db      ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1517 CATGACCTCTCGCTTTCTCAAGGTATACAGGCTGGTGCATGTCCTCATGATCAA 1576
QY 499 ProIleAlaSerLeuGlnLeuPheAsnArgThrIlePheGly 512
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1577 CCAAAAGCTGCATCTGAGATGCTAAGAGCTGGATGCAAGGC 1618
```

Search completed: July 29, 2004, 13:49:59

Job time : 723 secs